EXHIBIT 21

US006355432B1

(12) **United States Patent**
Fodor et al.

(10) **Patent No.: US 6,355,432 B1**
(45) **Date of Patent: Mar. 12, 2002**

(54) **PRODUCTS FOR DETECTING NUCLEIC ACIDS**

(75) Inventors: **Stephen P. A. Fodor**, Palo Alto; **Dennis W. Solas**, San Francisco; **William J. Dower**, Menlo Park, all of CA (US)

(73) Assignee: **Affymetrix Inc.**, Santa Clara, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/585,659**

(22) Filed: **Jun. 2, 2000**

Related U.S. Application Data

(63) Continuation of application No. 09/362,089, filed on Jul. 28, 1999, which is a division of application No. 09/056,927, filed on Apr. 8, 1998, now Pat. No. 6,197,506, which is a continuation of application No. 08/670,118, filed on Jun. 25, 1996, now Pat. No. 5,800,992, which is a division of application No. 08/168,904, filed on Dec. 15, 1993, now abandoned, which is a continuation of application No. 07/624,114, filed on Dec. 6, 1990, now abandoned, which is a continuation-in-part of application No. 07/492,462, filed on Mar. 7, 1990, now Pat. No. 5,143,854, which is a continuation-in-part of application No. 07/362,901, filed on Jun. 7, 1989, now abandoned.

(51) **Int. Cl.**[7] **C12Q 1/68**; G01N 33/53; C12M 1/34; C07H 21/04; A61K 38/00

(52) **U.S. Cl.** **435/6**; 435/7.1; 435/287.2; 435/288.3; 530/300; 530/350; 530/387.1; 536/23.1; 536/24.3

(58) **Field of Search** 536/23.1, 24.3; 435/6, 287.2, 288.3, 7.1; 530/300, 350, 387.1

(56) **References Cited**

U.S. PATENT DOCUMENTS

3,730,844 A	5/1973	Gilham et al.	195/103.5 R
3,849,137 A	11/1974	Barzynski et al.	96/97
3,862,056 A	1/1975	Hartman	252/511
3,939,350 A	2/1976	Kronick et al.	250/365
4,072,576 A	2/1978	Arwin et al.	195/103.5 R
4,121,222 A	10/1978	Diebold et al.	347/7
4,180,739 A	12/1979	Abu-Shumays	250/461 R
4,216,245 A	8/1980	Johnson	427/2.13
4,238,757 A	12/1980	Schenck	357/25
4,269,933 A	5/1981	Pazos	430/291
4,314,821 A	2/1982	Rice	23/230 B
4,327,073 A	4/1982	Huang	424/1
4,339,528 A	7/1982	Goldman	430/323
4,342,905 A	8/1982	Fujii et al.	250/201
4,373,071 A	2/1983	Itakura	525/375

(List continued on next page.)

FOREIGN PATENT DOCUMENTS

CA	1284931	6/1991
DE	2242394	3/1974
DE	3440141	5/1986
DE	3505287	3/1988
EP	046 083	2/1982
EP	088 636	9/1983
EP	103 197	3/1984
EP	127 438	12/1984
EP	063 810	3/1986
EP	174 879	3/1986
EP	194 132	9/1986
EP	228 075	7/1987
EP	245 662	11/1987
EP	268 237	5/1988
EP	130 523	6/1988

(List continued on next page.)

OTHER PUBLICATIONS

Brenner et al., "In vitro cloning of complex mixtures of DNA on microbeads: Physi–cal separation of differentially expressed cDNAs", PNAS, 02/2000, 97:665–1670.
Brenner et al., "Gene expression analysis by massively parallel signature sequen–cing (MPSS) on microbead arrays", Nature Biotechnol, 06/2000, 18:630–634.
Tyagi, "Taking a census of mRNA populations with micro-beads", Nature Biotechnol 06/2000, 18:597–598.
Miller et al. "Detection of bacteria by hybridization of rRNA with DNA–latex and immunodetection of hybrids" J Clin Microbiol 1988, 26:1271–1276.
Sequencing by Hybridization Workshop, listing of partici-pants and workshop presentation summaries, from work-shop held 11/19–20/91.
"A Sequencing Reality Check," *Science*, 242:1245 (1988).
"Affymax raises $25 million to develop high–speed drug discovery system," *Biotechnology News*, 10(3):7–8 (1990).
"Preparation of fluorescent–labeled DNA and its use as a probe in molecular hybridization," *Bioorg Khim*, 12(11):1508–1513 (1986).
Abbott et al., "Manipulation of the Wettability of Surfaces on the 0.1- to 1 –Micrometer Scale Through Micromachin-ing and Molcular Self–Assembly," *Science*, 257:1380–1382 (1992).
Adams et al., "Complementary DNA Sequencing: Expressed Sequence Tags and Human Genome Project," *Science*, 252(5013):1651–1656 (1991).
Adams et al., "Photolabile Chelators That "Cage" Calcium with Improved Speed of Release and Pre–Photolysis Affin-ity," *J. Gen. Physiol.*, p. 9a (12/86).
Adams et al., "Biologically Useful Chelators That Take Up Ca2+ upon Illumination," *J. Am. Chem. Soc.*, 111:7957–7968 (1989).
Ajayaghosh et al., "Solid–Phase Synthesis of N–Methyl– and N–Ethylamides of Peptides Using Photolytically Detachable ((3–Nitro–4((alkylamino)methyl)benzamido) methyl)polystyrene Resin," *J. Org. Chem.*, 55(9):2826–2829 (1990).

(List continued on next page.)

Primary Examiner—Stephanie W. Zitomer
(74) *Attorney, Agent, or Firm*—Morgan, Lewis & Bockius LLP

(57) **ABSTRACT**

The present invention provides methods and apparatus for sequencing, fingerprinting and mapping biological macromolecules, typically biological polymers. The meth-ods make use of a plurality of sequence specific recognition reagents which can also be used for classification of bio-logical samples, and to characterize their sources.

22 Claims, 2 Drawing Sheets

U.S. PATENT DOCUMENTS

4,395,486	A	7/1983	Wilson et al.	435/6
4,405,771	A	9/1983	Jagur	528/266
4,444,878	A	4/1984	Paulus	435/7
4,444,892	A	4/1984	Malmros	436/528
4,448,534	A	5/1984	Wertz et al.	356/435
4,458,066	A	7/1984	Caruthers et al.	536/27
4,483,920	A	11/1984	Gillespie et al.	435/6
4,500,707	A	2/1985	Caruthers et al.	536/27
4,500,919	A	2/1985	Schreiber	358/78
4,516,833	A	5/1985	Fusek	350/162.12
4,517,338	A	5/1985	Urdea et al.	525/54.11
4,533,682	A	8/1985	Tortorelio et al.	523/414
4,537,861	A	8/1985	Elings et al.	436/518
4,542,102	A	9/1985	Dattagupta et al.	435/6
4,555,490	A	11/1985	Merril	436/86
4,556,643	A	12/1985	Paau et al.	435/5
4,562,157	A	12/1985	Lowe et al.	435/291
4,563,419	A	1/1986	Ranki et al.	435/6
4,569,967	A	2/1986	Kornreich et al.	525/54.11
4,580,895	A	4/1986	Patel	356/39
4,584,277	A	4/1986	Ullman	436/501
4,588,682	A	5/1986	Groet et al.	435/6
4,591,570	A	5/1986	Chang	435/7.24
4,598,049	A	7/1986	Zelinka et al.	422/116
4,613,566	A	9/1986	Potter	435/6
4,624,915	A	11/1986	Schindler et al.	435/4
4,626,684	A	12/1986	Landa	250/328
4,631,211	A	12/1986	Houghten	428/35
4,637,861	A	1/1987	Krull et al.	204/1 T
4,656,127	A	4/1987	Mundy	435/6
4,670,380	A	6/1987	Dattagupta	435/6
4,677,054	A	6/1987	White et al.	435/6
4,681,859	A	7/1987	Kramer	436/501
4,683,195	A	7/1987	Mullis et al.	435/6
4,683,202	A	7/1987	Mullis	435/91
4,689,405	A	8/1987	Frank et al.	536/27
4,704,353	A	11/1987	Humphries et al.	435/4
4,711,955	A	12/1987	Ward et al.	536/29
4,713,326	A	12/1987	Dattagupta et al.	435/6
4,713,347	A	12/1987	Mitchell et al.	436/501
4,715,413	A	12/1987	Backlund et al.	141/94
4,716,106	A	12/1987	Chiswell	435/6
4,719,179	A	1/1988	Barany	435/172.1
4,719,615	A	1/1988	Feyrer et al.	369/284
4,722,906	A	2/1988	Guire	436/501
4,728,502	A	3/1988	Hamill	422/116
4,728,591	A	3/1988	Clark et al.	430/5
4,731,325	A	3/1988	Palva et al.	435/6
4,737,344	A	4/1988	Koizumi et al.	422/100
4,755,458	A	7/1988	Rabbani et al.	435/5
4,762,881	A	8/1988	Kauer	525/54.11
4,766,062	A	8/1988	Diamond et al.	435/6
4,767,700	A	8/1988	Wallace	435/6
4,777,019	A	10/1988	Dandekar	422/68
4,780,504	A	10/1988	Buendia et al.	525/54.11
4,786,170	A	11/1988	Groebler	356/318
4,786,684	A	11/1988	Glass	525/54.1
4,794,150	A	12/1988	Steel	525/54.11
4,808,508	A	2/1989	Platzer	430/143
4,810,869	A	3/1989	Yabe et al.	250/201
4,811,062	A	3/1989	Tabata et al.	356/152
4,811,218	A	3/1989	Hunkapiller et al.	204/461
4,812,512	A	3/1989	Buendia et al.	525/54.11
4,820,630	A	4/1989	Taub	435/5
4,822,566	A	4/1989	Newman	422/68
4,833,092	A	5/1989	Geysen	436/501
4,844,617	A	7/1989	Kelderman et al.	356/372
4,846,552	A	7/1989	Veldkamp et al.	350/162.2
4,849,513	A	7/1989	Smith et al.	536/27
4,855,225	A	8/1989	Fung et al.	435/6
4,865,990	A	9/1989	Stead et al.	435/803
4,868,103	A	9/1989	Stavrianopoulos et al.	435/5
4,874,500	A	10/1989	Madou et al.	204/412
4,877,745	A	10/1989	Hayes et al.	436/166
4,886,741	A	12/1989	Schwartz	435/5
4,888,278	A	12/1989	Singer et al.	435/6
4,921,805	A	5/1990	Gebeyehu et al.	435/270
4,923,901	A	5/1990	Koester et al.	521/53
4,925,785	A	5/1990	Wang et al.	435/6
4,931,384	A	6/1990	Layton et al.	435/7.31
4,946,942	A	8/1990	Fuller et al.	530/335
4,965,188	A	10/1990	Mullis et al.	435/6
4,973,493	A	11/1990	Guire	427/2
4,979,959	A	12/1990	Guire	623/66
4,981,783	A	1/1991	Augenlicht	435/6
4,981,985	A	1/1991	Kaplan et al.	556/50
4,984,100	A	1/1991	Takayama et al.	360/49
4,987,065	A	1/1991	Stavrianopoulos et al.	435/5
4,988,617	A	1/1991	Landegren et al.	435/6
4,992,383	A	2/1991	Farnsworth	436/89
4,994,373	A	2/1991	Stavrianopoulos et al.	435/6
5,002,867	A	3/1991	Macevicz	435/6
5,006,464	A	4/1991	Chu et al.	435/7.1
5,011,770	A	4/1991	Kung et al.	435/6
5,013,669	A	5/1991	Peters, Jr. et al.	436/518
5,021,550	A	6/1991	Zieger	530/334
5,026,773	A	6/1991	Steel	525/54.11
5,026,840	A	6/1991	Dattagupta et al.	536/27
5,028,525	A	7/1991	Gray et al.	435/6
5,028,545	A	7/1991	Soini	436/501
5,037,882	A	8/1991	Steel	525/54.11
5,043,265	A	8/1991	Tanke et al.	435/6
5,047,524	A	9/1991	Andrus et al.	536/27
5,064,754	A	11/1991	Mills	435/6
5,075,077	A	12/1991	Durley	
5,077,085	A	12/1991	Schnur et al.	427/98
5,077,210	A	12/1991	Eigler et al.	435/176
5,079,600	A	1/1992	Schnur et al.	357/4
5,081,584	A	1/1992	Omichinski et al.	364/497
5,082,830	A	1/1992	Brakel et al.	514/44
5,091,652	A	2/1992	Mathies et al.	250/458.1
5,096,807	A	3/1992	Leaback	435/6
5,100,626	A	3/1992	Levin	422/100
5,100,777	A	3/1992	Chang	435/7.24
5,112,962	A	5/1992	Letsinger et al.	536/27
5,141,813	A	8/1992	Nelson	428/402
5,143,854	A	9/1992	Pirrung et al.	436/518
5,149,625	A	9/1992	Church et al.	435/6
5,153,319	A	10/1992	Caruthers et al.	536/27
5,164,319	A	11/1992	Hafeman et al.	435/287.1
5,171,695	A	12/1992	Ekins	436/501
5,188,963	A	2/1993	Stapleton	435/288.3
5,192,980	A	3/1993	Dixon et al.	356/326
5,200,051	A	4/1993	Cozzette et al.	204/403
5,202,231	A	4/1993	Drmanac et al.	435/6
5,206,137	A	4/1993	Ip et al.	435/6
5,215,882	A	6/1993	Bahl et al.	435/6
5,215,889	A	6/1993	Schultz	435/41
5,219,726	A	6/1993	Evans	435/6
5,225,326	A	7/1993	Bresser et al.	435/6
5,232,829	A	8/1993	Longiaru et al.	435/6
5,235,028	A	8/1993	Barany et al.	528/335
5,242,974	A	9/1993	Holmes	525/54.11
5,252,743	A	10/1993	Barrett et al.	548/303.7
5,256,549	A	10/1993	Urdea et al.	435/91
5,258,506	A	11/1993	Urdea et al.	536/23.1
5,306,641	A	4/1994	Saccocio	436/85
5,310,893	A	5/1994	Erlich et al.	536/24.31
5,324,633	A	6/1994	Fodor et al.	435/6
5,328,824	A	7/1994	Ward et al.	435/6
5,348,855	A	9/1994	Dattagupta et al.	435/6

5,384,261 A	1/1995	Winkler et al.	436/518
5,405,783 A	4/1995	Pirrung et al.	436/518
5,424,186 A	6/1995	Fodor et al.	435/6
5,424,188 A	6/1995	Schneider et al.	435/6
5,432,099 A	7/1995	Ekins	436/518
5,436,327 A	7/1995	Southern et al.	536/25.34
5,445,934 A	8/1995	Fodor et al.	435/6
5,447,841 A	9/1995	Gray et al.	435/6
5,451,505 A	9/1995	Dollinger	
5,474,796 A	12/1995	Brennan	427/2.13
5,486,452 A	1/1996	Gordon et al.	435/5
5,489,507 A	2/1996	Chehab	435/6
5,489,678 A	2/1996	Fodor et al.	536/22.1
5,492,806 A	2/1996	Drmanac et al.	435/5
5,494,810 A	2/1996	Barany et al.	435/91.52
5,510,270 A	4/1996	Fodor et al.	436/518
5,525,464 A	6/1996	Drmanac et al.	435/6
5,527,681 A	6/1996	Holmes	435/6
5,552,270 A	9/1996	Khrapko et al.	435/6
5,556,961 A	9/1996	Foote et al.	536/27.1
5,561,071 A	10/1996	Hollenberg et al.	437/1
5,565,324 A	10/1996	Still	
5,567,809 A	10/1996	Mandecki	
5,569,584 A	10/1996	Augenlicht	435/6
5,571,639 A	11/1996	Hubbell et al.	430/5
5,573,905 A	11/1996	Lerner	
5,593,839 A	1/1997	Hubbell et al.	435/6
5,599,720 A	2/1997	Ekins	436/501
5,604,097 A	2/1997	Brenner	
5,604,099 A	2/1997	Erlich et al.	435/6
5,635,400 A	6/1997	Brenner	
5,641,634 A	6/1997	Mandecki	
5,643,728 A	7/1997	Slater et al.	435/6
5,653,939 A	8/1997	Hollis et al.	422/50
5,654,413 A	8/1997	Brenner	
5,667,667 A	9/1997	Southern	205/687
5,667,972 A	9/1997	Drmanac et al.	435/6
5,690,894 A	11/1997	Pinkel	
5,695,940 A	12/1997	Drmanac et al.	435/6
5,698,393 A	12/1997	Macioszek et al.	435/5
5,700,637 A	12/1997	Southern	435/6
5,707,806 A	1/1998	Shuber	435/6
5,744,305 A	4/1998	Fodor et al.	435/6
5,751,629 A	5/1998	Nova	
5,770,367 A	6/1998	Southern	
5,776,737 A	7/1998	Dunn	435/91.1
5,777,888 A	7/1998	Rine et al.	364/496
5,800,992 A	9/1998	Fodor et al.	435/6
5,804,563 A	9/1998	Still	
5,807,522 A	9/1998	Brown	
5,807,683 A	9/1998	Brenner	
5,830,645 A	11/1998	Pinkel et al.	435/6
5,843,767 A	12/1998	Beattie	435/287.1
5,846,708 A	12/1998	Hollis et al.	435/6
5,846,719 A	12/1998	Brenner	
5,863,722 A	1/1999	Brenner	
5,869,237 A	2/1999	Ward et al.	435/6
5,871,697 A	2/1999	Rothberg et al.	422/68.1
5,972,619 A	10/1999	Drmanac et al.	435/6
6,018,041 A	1/2000	Drmanac et al.	536/24.3
6,023,540 A	2/2000	Walt	
6,025,136 A	2/2000	Drmanac et al.	435/6
6,040,166 A	3/2000	Erlich et al.	435/194
6,054,270 A	4/2000	Southern	435/6
6,060,240 A	5/2000	Kamb	

FOREIGN PATENT DOCUMENTS

EP	281 927	9/1988
EP	228 310	10/1988
EP	288 310	10/1988
EP	304 202	2/1989
EP	307 476	3/1989
EP	319 012	6/1989
EP	328 256	8/1989
EP	333 561	9/1989
EP	337 498	10/1989
EP	386 229	4/1990
EP	373 203	6/1990
EP	392 546	10/1990
EP	142 299	12/1990
EP	173 339	1/1992
EP	171 150	3/1992
EP	237 362	3/1992
EP	185 547	6/1992
EP	260 634	6/1992
EP	232 967	4/1993
EP	235 726	5/1993
EP	476 014	8/1994
EP	225 807	10/1994
EP	717 113	6/1996
EP	721 016	7/1996
EP	535 242	9/1997
EP	848 067	6/1998
EP	619 321	1/1999
FR	2559783	3/1988
GB	2 129 551	5/1984
GB	2156074	3/1988
GB	2196476	4/1988
GB	8810400.5	5/1988
GB	2233654	1/1991
GB	2248840	9/1992
JP	49-110601	10/1974
JP	60-248669	12/1985
JP	63-084499	4/1988
JP	63-223557	9/1988
JP	1-233447	9/1989
NO	P 913186	8/1991
WO	WO 84/03151	8/1984
WO	WO 84/03564	9/1984
WO	WO 85/01051	3/1985
WO	WO 86/00991	2/1986
WO	WO 86/06487	11/1986
WO	WO 88/01058	2/1988
WO	WO 88/04777	6/1988
WO	WO 89/05616	6/1989
WO	WO 89/08834	9/1989
WO	WO 89/10977	11/1989
WO	WO 89/11548	11/1989
WO	WO 89/12819	12/1989
WO	WO 90/00626	1/1990
WO	WO 90/00887	2/1990
WO	WO 90/15070	2/1990
WO	WO 90/03382	4/1990
WO	WO 90/04652	5/1990
WO	WO 90/05789	5/1990
WO	WO 90/07582	7/1990
WO	WO 91/00868	1/1991
WO	WO 91/04266	4/1991
WO	WO 91/07087	5/1991
WO	WO 92/16655	1/1992
WO	WO 92/10092	6/1992
WO	WO 92/10588	6/1992
WO	WO 93/02992	2/1993
WO	WO 93/09668	5/1993
WO	WO 88/01302	6/1993
WO	WO 93/11262	6/1993
WO	WO 93/17126	9/1993
WO	WO 93/22456	11/1993
WO	WO 93/22480	11/1993
WO	WO 95/00530	1/1995
WO	WO 95/11995	5/1995
WO	WO 95/33846	12/1995

WO	WO 96/23078	8/1996
WO	WO 97/10365	3/1997
WO	WO 97/17317	5/1997
WO	WO 97/19410	5/1997
WO	WO 97/27317	7/1997
WO	WO 97/29212	8/1997
WO	WO 97/31256	8/1997
WO	WO 97/45559	12/1997
WO	WO 98/03673	1/1998
WO	WO 98/31836	7/1998
WO	WO 99/60007	11/1999
YU	P-570/87	4/1987
YU	18617/87	9/1987

OTHER PUBLICATIONS

Ajayaghosh et al., "Solid–phase synthesis of C–terminal peptide amides using a photoremovable α–methyphenacylamido anchoring linkage," *Proc. Ind. Natl. Sci* (*Chem.Sci.*), 100(5):389–396 (1988).

Ajayaghosh et al., "Polymer–supported Solid–phase Synthesis of C–Terminal Peptide N–Methylamides Using a Modified Photoremovable 3–Nitro–4–N–methylaminomethylpolystyrene Support," *Ind.J.Chem.*, 27B:1004–1008 (1988).

Ajayaghosh et al., "Polymer–Supported Synthesis of Protected Peptide Segments on a Photosensitive o–Nitro (α–Methyl)Bromobenzyl Resin," *Tetrahedron*, 44(21): 6661–6666 (1988).

Amit et al., "Photosensitive Protecting Groups of Amino Sugars and Their Use in Glycoside Synthesis. 2–Nitrobenzyloxycarbonylamino and 6–Nitroveratryloxycarbonylamino Derivatives," *J.Org.Chem*, 39(2):192–196 (1974).

Amit et al., "Photosensitive Protecting Groups—A Review," *Israel J. Chem.*, 12(1–2):103–113 (1974).

Anand et al., "A 3.5 genome equivalent multi access YAC library: construction, characterisation, screeing and storage," *Nuc. Acids Res.*, 18(8):1951–1956 (1990).

Anderson et al., "Quantitative Filter Hybridisation," chapter 3 from Nucleic Acid Hybridization a practical approach, pp. 73–111, Hames et al., eds., IRL Press (1985).

Applied Biosystems, Model 431A Peptide Synthesizer User's manual, Sections 2 and 6, (Aug. 15, 1989).

Anold et al., "A Novel Universal Support for DNA & RNA Synthesis," abstract from *Federation Proceedings*, 43(7):abstract No. 3669 (1984).

Atherton et al., Solid Phase Peptide Synthesis: A Practical Approach, IRL Press, (1989), tbl. of cont., pp. vii–ix.

Augenlicht et al.,"Cloning and Screening of Sequences Expressed in a Mouse Colon Tumor," *Cancer Research*, 42:1088–1093 (1982).

Augenlicht et al., "Expression of Cloned Sequences in Biopsies of Human Colonic Tissue and in Colonic Carcinoma Cells Induced to Differentiate in Vitro," *Cancer Res.*, 47:6017–6021 (1987).

Bains, W., "Hybridization Methods for DNA Sequencing," *Genomics*, 11(2):294–301 (1991).

Bains et al., "A Novel Method for Nucleic Acid Sequence Determination," *J.Theor.Biol.*, 135:303–307 (1988).

Bains, W., "Alternative Routes Through the Genome," *Biotechnology*, 8:1251–1256 (1988).

Balachander et al., "Functionalized Siloxy–Anchored Monolayers with Exposed Amino, Azido, Bromo, or Cyano Groups," *Tetrahed. Ltrs.*, 29(44):5593–5594 (1988).

Baldwin et al., "New Photolabile Phosphate Protecting Groups," *Tetrahed.*, 46(19):6879–6884 (1990).

Bannwarth et al., "Laboratory Methods, A System for the Simultaneous Chemical synthesis of Different DNA Fragments on Solid Support," *DNA*, 5(5):413–419 (1986).

Bannwarth, W., "Gene Technology: a Challenge for a Chemist," *Chimia*, 41(9):302–317 (1987).

Barany, F., "Genetic disease detection and DNA amplification using cloned thermostable ligase," *PNAS*, 88:189–193 (1991).

Barltrop et al., "Photosensitive Protective Groups," *Chemical Communications*, pp. 822–823 (1966).

Barinaga, M., "Will 'DNA Chip' Speed Genome Initiative," *Science*, 253:1489 (1985).

Bart et al., "Microfabricated Electrohydrodynamic Pumps," *Sensors and Actuators*, A21:–A23:193–197 (1990).

Bartsh et al., "Cloning of mRNA sequences from the human colon: Preliminary characterisation of defined mRNAs in normal and neoplastic tissues," *Br.J.Can.*, 54:791–798 (1986).

Baum, R., "Fledgling firm targets drug discovery process," *Chem. Eng. News*, p. 10–11 (1990).

Beltz et al., "Isolation of Multigene Families and Determination of Homologies by Filter Hybridization Methods," *Methods in Enzymology*, 100:266–285 (1983).

Benschop, Chem. Abstracts 114(26):256643 (1991).

Bhatia et al., "New Approach To Producing Patterned Biomolecular Assemblies," *J. American Chemical Society*, 114:4432–4433 (1992).

Biorad Chromatography Electrophoresis Immunochemistry Molecular Biology HPLC catalog M 1987 pp. 182.

Blawas et al., "Step–and–Repeat Photopatterning of Protein Features Using Caged–Biotin–BSA: Characterization and Resolution," *Langmuir*, 14(15):4243–4250 (1998).

Blawas, A.S., "Photopatterning of Protein Features using Caged–biotin–Bovine Serum Albumin," dissertation for Ph.D at Duke University in 1998.

Bos et al., "Amino–acid substirutions at codon 13 of the N–ras oncogene in human acute myeloid leukaemia," *Nature*, 315:726–730 (1985).

Boyle et al., "Differential distribution of long and short interspersed element sequences in the mouse genome: Chromosome karyotyping by fluorescence in situ hybridization," *PNAS*, 87:7757–7761 (1990).

Brock et al., "Rapid fluorescence detection of in situ hybridization with biotinylated bovine herpesvirus–1 DNA probes," *J. Veterinary Diagnostic Invest.*, 1:34–38 (1989).

Burgi et al., "Optimization in Sample Stacking for High–Performance Capillary Electrophoresis," *Anal. Chem.*, 63:2042–2047 (1991).

Cameron et al., "Photogeneration of Organic Bases from o–Nitrobenzyl–Derived Carbamates," *J. Am. Chem. Soc.*, 113:4303–4313 (1991).

Carrano et al., "A High–Resolution, Fluorescence–Based, Semiautomated Method for DNA Fingerprinting," *Genomics*, 4:129–136 (1989).

Caruthers, M.H., "Gene Synthesis Machines: DNA Chemistry and Its Uses," *Science*, 230:281–285 (1985).

Chatterjee et al., "Inducible Alkylation of DNA Using an Oligonucleotide–Quinone Conjugates," *Am. J. Chem. Soc.*, 112:6397–6399 (1990).

Chee et al., "Accessing Genetic Information with High–Density DNA Arrays," *Science*, 274:610–614 (1996).

Chehab et al., "Detection of sicle cell anaemia mutation by colour DNA amplification," *Lancet*, 335:15–17 (1990).

Chehab et al., "Detection of specific DNA sequences by fluorescence amplification: A color complementation assay," *PNAS*, 86:9178–9182 (1989).

Chetverin et al., "Oligonucleotide Arrays: New Concepts and Possibilities," *Biotechnology*, 12:1093–1099 (1994).

Church et al., "Multiplex DNA sequencing," *Science*, 240:185–188 (1988).

Church et al., "Genomic sequencing," *PNAS*, 81:1991–1995 (1984).

Clevite Corp., Piezoelectric Technology, Data for Engineers.

Corbett et al., "Reaction of Nitroso Aromatics with Glyoxylic Acid. A New Path to Hydroxamic Acids," *J. Org. Chem.*, 45:2834–2839 (1980).

Coulson et al., "Toward a physical map of the genome of the nematode *Caenorhabditis elegans*," *PNAS*, 83:7821–7825 (1986).

Craig et al., "Ordering of cosmid clones covering the Herpes simplex virus type 1 (HSV–1) genome: a test case for fingerprinting the hybridization," *Nuc. Acid. Res.*, 18(9):2653–2660 (1990).

Cummings et al., "Photoactivable Fluorophores. 1. Synthesis and Photoactivation of o–Nitrobenzyl–Quenched Fluorescent Carbamates," *Tetrahedron Letters*, 29(1):65–68 (1988).

Dattagupta et al., "Rapid identification of Microorganisms by Nucleic Acid Hybridization after Labeling the Test Sample," *Anal. Biochem.*, 177:85–89 (1989).

Dattagupta et al., "Nucleic Acid Hybridization: a Rapid Method for the Diagnosis of Infectious Diseases," *Perspectives in Antiinfective Therapy*, eds. Jackson et al., pp. 241–247 (1988).

Dower et al., "The Search for Molecular Diversity (II): Recombinant and Synthetic Randomized Peptide Libraries," *Ann. Rep. Med. Chem.*, 26:271–280 (1991).

Diggelmann, "Investigating the VLSIPS synthesis process," Sep. 9, 1994.

Di Mauro et al., "DNA Technology in Chip Construction," *Adv. Mater.*, 5(5):384–386 (1993).

Drmanac et al., "An Algorithm for the DNA Sequence Generation from k–Tuple Word Contents of the Minimal Number of Random Fragments," *J. Biomol.Struct.Dyn.*, 8(5):1085–1102 (1991).

Drmanac et al., "Partial Sequencing by Oligo–Hybridization Concept and Applications in Genome Analysis," 1st Int. Conf. Electrophor., Supercomp., Hum. Genome pp. 60–74 (1990).

Drmanac et al., "Sequencing by Oligonucleotide Hybridization: A Promising Framework in Decoding of the Genome Program?," 1st Int. Conf. Electrophor., Supercomp., Hum. Genome pp. 47–59 (1990).

Drmanac et al., "Laboratory Methods, Reliable Hybridization: theory of the Method," *Genomics*, 4:114–128 (1989).

Dramanac et al., "Sequencing of Megabase Plus DNA by Hyrbridization: Theory of the Method," abstract of presentation given at Cold Spring Harbor Symposium on Genome Mapping and Sequencing, Apr. 27, 1988 thru May 1, 1988.

Dulcey et al., "Deep UV Photochemistry of Chemisorbed Monolayers: Patterned Coplanar Molecular Assemblies," *Science*, 252:551–554 (1991).

Duncan et al., "Affinity Chromatography of a Sequence–Specific DNA Binding Protein Using Teflon–Linked Oligonucleotides," *Analytical Biochemistry*, 169:104–108 (1988).

Effenhauser et al., "Glass Chips for High–speed Capillary Electrophoresis Separations with Submicrometer Plate Heights," Anal. Chem., 65:2637–2642 (1993).

Effenhauser et al., "High–Speed Separation of Antisense Oligonucleotides on a Micromachined Capillary Electrophoresis Device," *Anal. Chem.*, 66:2949–2953 (1994).

Ekins et al., "High Specific Activity Chemilminescent and Fluorescent Markers: their Potential Application to High Sensitivity and 'Multi–analyte' Immunoassays," *J. Biolumi-nescence Chemiluminescence*, 4:59–78 (1989).

Ekins et al., "Development of Microspot Multi–Analyte Ratiometric Immunoassay Using dual Fluorescent–Labelled Antibodies," *Anal. Chemica Acta*, 227:73–96 (1989).

Ekins et al., "Multianalyte Microspot Immunoassay–Microanalytical 'Compact Disk' of the Future," *Clin. Chem.*, 37(11):1955–1967 (1991).

Ekins, R.P., "Multi–Analyte immunoassay*," *J. Pharmaceut. Biomedical Analysis*, 7(2):155–168 (1989).

Ekins et al., "Fluorescence Spectroscopy and its Application to a New Generation of High Sensitivity, Multi–Microspot, Multianalyte, Immunoassay," *Clin. Chim. Acta*, 194:91–114 (1990).

Elder, J.K., "Analysis of DNA Oligonucleotide Hybridization Data by Maximum Entropy," *Maximum Entropy and Bayesian Methods*, eds. Mohammad–Djafari and Demoment, Kluwer, Dordrecht, pp. 363–371 (1992).

Ellis, R.W., "The Applications of Synthetic Oligonucleotides to Molecular Biology," *Pharmaceutical Research*, 3(4):195–207 (1986).

Evans et al., "Microfabrication for Automation of Molecular processes im Human Genome Analysis," *Clin. Chem.*, 41(11):1681 (1995).

Evans et al., "Physical mapping of complex genomes by cosmid multiplex analysis," *PNAS*, 86:5030–5034 (1989).

Ezaki et al., "Small–Scale DNA Preparation for Rapid Genetic Identification of Campylobacter Species without Radioisotope," *Microbiol. Immunology*, 32(2):141–150 (1988).

Fan et al., "Mapping small DNA sequences by fluorescence in situ hybridization directly on banded metaphase chromosomes," PNAS, 87(16):6223–6227 (1990).

Fan et al., "Micromachining of Capillary Electrophoresis Injectors and Separators on Glass Chips and Evaluation of Flow at Capillary Intersections," Anal. Chem., 66:177–184 (1994).

Feinberg et al., Addendum to "A technique for Radiolabeling DNA Restriction Endonuclease Fragments to High Specific Activity," *Anal. Biochem.*, 137:266–267 (1984).

Fettinger et al., "Stacked modules for micro flow systems in chemical analysis: concept and studies using an enlarged model," *Sensors and Actuators*, B17:19–25 (1993).

Flanders et al., "A new interferometric alignment technique," *App. Phys. Ltrs.*, 31(7):426–429 (1977).

Fodor et al., "Multiplexed biochemical assays with biological chips," *Nature*, 364:555–556 (1993).

Fodor et al., "Light–directed, Spatially Addressable Parallel Chemical Synthesis," *Science*, 251:767–773 (1991).

Forman et al., "Thermodynamics of Duplex Formation and Mismatch Discrimination on Photolithographically Synthesized Oligonucleotide Arrays," chapter 13pgs. 206–228 from *Molecular Modeling of Nucleic Acids*, ACS Symposium Series 682, 4/13–17/97, Leontis et al., eds.

Frank et al., "Simultaneous Multiple Peptide Synthesis Under Continuous flow Conditions on Cellulose Paper Discs as Segmental Solid Supports," *Tetrahedron*, 44(19):6031–6040 (1988).

Frank et al., "Automation of DNA Sequencing Reactions and Related Techniques: A Workstation for Micromanipulation of Liquids," *Bio/Technology*, 6:1211–1212 (1988).

Frank et al., "Simultaneous Synthesis and Biological Applications of DNA Fragments: An Efficient and Complete Methodology," *Methods in Enzymology*, 154:221–250 (1987).

Fuhr et al., "Travelling wave–driven microfabricated electrohydrodynamic pumps for liquids," *J. Micromech. Microeng.*, 4:217–226 (1994).

Fuller et al., "Urethane–Protected Amino Acid N–Carboxy Anhydrides and Their Use in Peptide Synthesis," *J. Amer. Chem. Soc.*, 112(20):7414–7416 (1990).

Furka et al., "General method for rapid synthesis of multicomponent peptide mixtures," *Int. J. Peptide Protein Res.*, 37:487–493 (1991).

Furka et al., "Cornucopia of Peptides by Synthesis," 14th Int.Congress of Biochem. abst.# FR:013, 7/10–15/88 Prague, Czechoslovakia.

Furka et al., "More Peptides by Less Labour," abst. 288, Int. Symp. Med. Chem., Budapest Hungary.

Gait, eds., pp. 1–115 from *Oligonucleotide Synthesis: A Practical Approach*, IRL Press, (1984).

Gazard et al., "Lithographic Technique Using Radiation–Induced Grafting of Acrylic Acid into Poly(Methyl Methacrylate) Films," *Polymer Engineering and Science*, 20(16):1069–1072 (1980).

Gergen et al., "Filter replicas and permanent collections of recombinant DNA plasmids," *Nuc.Acids Res.*, 7(8):2115–2137 (1979).

Getzoff et al., "Mechanisms of Antibody Binding to a Protein," *Science*, 235:1191–1196 (1987).

Geysen et al., "Strategies for epitope analysis using peptide synthesis," *J. Immunol. Meth.*, 102:259–274 (1987).

Geysen et al., "Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid," *PNAS*, 81:3998–4002 (1984).

Geysen et al., "A synthetic strategy for epitope mapping," from Peptides:Chem. & Biol., Proc. of 10th Am. Peptide Symp., 5/23–28/87, pp. 519–523, (1987).

Geysen, "Antigen–antibody interactions at the molecular level: adventures in peptide synthesis," *Immunol. Today*, 6(12):364–369 (1985).

Geysen et al., "Cognitive Features of Continuous Antigenic Determinants," from Synthetic Peptides: Approaches to Biological Probes, pp. 19–30, (1989).

Geysen et al., "Chemistry of Antibody Binding to a Protein," *Science*, 235:1184–1190 (1987).

Geysen et al., "The delineation of peptides able to mimic assembled epitopesm" 1986 CIBA Symp., pp. 130–149.

Geysen et al., "Cognitive Features of Continuous Antigenic Determinants," *Mol. Recognit.*, 1(1):1–10 (1988).

Geysen et al., "A Prio Ri Delineation of a Peptide Which Mimics A Discontinuous Antigenic Determinant," *Mol. Immunol.*, 23(7):709–715 (1986).

Ghosh et al., "Covalent attachment of oligonucleotides to solid supports," *Nuc. Acids Res.*, 15(13):5353–5373 (1987).

Gilon et al., "Backbone Cyclization: A New Method for Conferring Conformational Constraint on Peptides," *Biopolymers*, 31(6):745–750 (1991).

Gingeras et al., "Hybridization properties of immobilized nucleic acids," *Nuc. Acids Res.*, 15(13):5373–5390 (87).

Gummerlock et al., "RAS Enzyme–Linked Immunoblot Assay Discriminates p21 Species: A Technique to Dissect Gene Family Expression," *Anal. Biochem.*, 180:158–168 (1989).

Gurney et al., "Activation of a potassium current by rapid photochemically generated step increases of intracellular calcium in rat sympathetic neurons," *PNAS*, 84:3496–3500 (1987).

Haase et al., "Detection of Two Viral Genomes in Single Cells by Double–Label Hybridization in Situ and Color Microradioautography," *Science*, 227:189–192 (1985).

Hacia, et al., "Two color hybridization analysis using high density oligonucleotide arrays and energy transfer dyes," *Nuc. Acids Res.*, 26(16):3865–3866 (1998).

Hack, M.L., "Conics Formed to Make Fluid & Industrial Gas Micromachines," *Genetic Engineering News*, 15(18):1, 29 (1995).

Hagedorn et al., "Pumping of Water Solutions in Microfabricated Electrohydrodynamic Systems," from Micro Electro Mechanical Systems conference in Travemunde Germany (1992).

Hames et al., *Nuclear acid hybridization, a practical approach*, cover page and table of contents (1985).

Hanahan et al., "Plasmid Screening at High Colony Density," *Meth. Enzymology*, 100:333–342 (1983).

Hanahan et al., "Plasmid screening at high colony density," *Gene*, 10:63–67 (1980).

Haridasan et al., "Peptide Synthesis using Photolytically Cleavable 2–Nitrobenzyloxycarbonyl Protecting Group," *Proc. Indian Natn. Sci. Adad.*, 53A(6):717–728 (1987).

Harrison et al., "Capillary Electrophoresis and Sample Injection Systems Integrated on a Planar Glass Chip," *Anal. Chem.*, 64:1926–1932 (1992).

Harrison et al., "Micromachining a Minaturized Capillary Electrophoresis–Based Chemical Analysis System on a Chip," *Science*, 261:895–897 (1993).

Harrison et al., "Towards minaturized electrophoresis and chemical analysis systems on silicon: an alternative to chemical sensors*," *Sensors and Actuators*, B10:107–116 (1993).

Harrison et al., "Rapid separation of fluorescein derivatives using a micromachined capillary electrophoresis system," *Analytica Chemica Acta*, 283:361–366 (1993).

Hellberg et al., "Minimum analogue peptide sets (MAPS) for quantitative structure–activity relationships," *Int. J. Peptide Protein Res.*, 37:414–424 (1991).

Hilser et al., "Protein and peptide mobility in capillary zone electrophoresis, A comparison of existing models and further analysis," *J. Chromatography*, 630:329–336 (1993).

Ho et al., "Highly Stable Biosensor Using an Artificial Enzyme," *Anal.Chem.*, 59:536–537 (1987).

Hochgeschwender et al., "Preferential expression of a defined T–cell receptor β–chain gene in hapten–specific cytotoxic T–cell clones," *Nature*, 322:376–378 (1986).

Hodgson, J., "Assays A La Photolithography," *Biotech.*, 9:419 (1991).

Hodgson et al., "Hybridization probe size control: optimized 'oligolabelling'," *Nuc.Acids Res.*, 15(15):6295 (1987).

Hopman et al., "Bi–color detection of two target DNAs by non–radioactive in situ hybridization*," *Histochem.*, 85:1–4 (1986).

Iwamura et al., "1–Pyrenylmethyl Esters, Photolabile Protecting Groups for Carboxlic Acids," *Tetrahedron Ltrs.*, 28(6):679–682 (1987).

Iwamura et al., "1–(α–Diazobenzyl)pyrene: A Reagent for Photolabile and Fluorescent Protection of Carboxyl Groups of Amino Acids and Peptides," *Synlett*, p. 35–36 (1991).

Jacobson et al., "Effects of Injection Schemes and Column Geometry on the Performance of Microchip Electrophoresis Devices," *Anal. Chem.*, 66:1107–1113 (1994).

Jacobsen et al., "Open Channel Electrochromatography on a Microchip," *Anal. chem.*, 66:2369–2373 (1994).

Jacobson et al., "Microchip Capillary Electrophoresis with an Integrated Postcolumn Reactor" *Anal. Chem.*, 66:3472–3476 (1994).

Jacobson et al., "Precolumn Reactions with Electrophoretic Analysis Integrated on a Microchip," *Anal. Chem.*, 66:4127–4132 (1994).

Jacobson et al., "Microfabricated chemical measurement systems," *Nature Medicine*, 1(10):1093–1096 (1995).

Jacobsen et al., "Fused Quartz Substrates for Microchip Electrophoresis," *Anal. chem.*, 67:2059–2063 (1995).

Jacobson et al., "High–Speed Separtions on a Microchip," *Anal. Chem.*, 66:1114–1118 (1994).

Jacobson et al., "Microchip electrophoresis with sample stacking," *Electrophoresis*, 16:481–486 (1995).

Jayakumari, "Peptide synthesis in a triphasic medium catalysed by papain immobilized on a crosslinked polystyrene support," *Indian J. Chemistry*, 29B:514–517 (1990).

Jovin et al., "Luminescence Digital Imaging Microscopy," *Ann. Rev. Biophys. Biophys. Chem.*, 18:271–308 (1989).

Kafatos et al., "Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure," *Nuc. Acids Res.*, 7(6):1541–1553 (1979).

Kaiser et al., "Peptide and Protein Synthesis by Segment Synthesis–Condensation," *Science*, 243:187–192 (1989).

Kaplan et al., "Photolabile chelators for the rapid photorelease of divalent cations," *PNAS*, 85:6571–6575 (1988).

Karube, "Micro–biosensors based on silicon fabrication technology," chapter 25 from Biosensors:Fundamentals and Applications, Turner et al., eds., Oxford Publ., 1987, pp. 471–480 (1987).

Kates et al., "A Novel, Convenient, Three–dimensional Orthogonal Strategy for Solid–Phase Synthesis of Cyclic Peptides 1–3," *Tetrahed. Letters*, 34(10):1549–1552 (1993).

Kerkof et al., "A Procedure for Making Simultaneous Determinations of the Relative Levels of Gene Transcripts in Tissues or Cells," *Anal. Biochem.*, 188:349–355 (1990).

Khrapko et al., "An Oligonucleotide hybridization approach to DNA sequencing," *FEBS Lett.*, 256(1,2):118–122 (1989).

Khrapko et al., "A method for DNA sequencing by hybridization with oligonucleotide matrix," *DNA Seq. Map.*, 1:375–388 (1991).

Kidd et al., "α_1–Antitrypsin deficiency detection by direct analysis of the mutation in the gene," *Nature*, 304:230–234 (1983).

Kievits et al., "Rapid subchromosomal localization of cosmids by nonradioactive in situ hybridization," *Cytogenetics Cell Genetics*, 53(2–3):134–136 (1990).

Kimura et al., "An Immobilized Enzyme Membrane Fabrication Method using an Ink Jet Nozzle," *Biosensors*, 4:41–52 (1988).

Kimura et al., "An Integrated SOS/FET Multi–Biosensor," *Sensors & Actuators*, 9:373–387 (1986).

Kitazawa et al., "In situ DNA–RNA hybridization using in vivo bromodeoxyuridine–labeled DNA probe," *Histochemistry*, 92:195–199 (1989).

Kleinfeld et al., "Controlled Outgrowth of Dissociated Neurons on Patterned Substrates," *J. Neurosci.*, 8(11):4098–4120 (1988).

Knight, P., "Materials and Methods/Microsequencers for Proteins and Oligosaccharides," *Bio/Tech.*, 7:1075–76 (1989).

Kohara et al., "The Physical Map of the Whole *E. coli* Chromosome: Application of a New Strategy for Rapid Analysis and Sorting of a Large Genomic Library," *Cell*, 50:495–508 (1987).

Krile et al., "Multiplex holography with chirp–modulated binary phase–coded reference–beam masks," *Applied Opt.*, 18(1):52–56 (1979).

Labat, I., "Subfragments as an informative characteristic of the DNA molecule—computer simulation," research report submitted to the University of Belgrade College of Natural Sciences and Mathematics, (1988).

Lander et al., "Genomic Mapping by Fingerprinting Randon Clones: A Mathematical Analysis," *Genomics*, 2:231–239 (1988).

Lainer et al., "Human Lymphocyte Subpopulations Identified by Using Three–Color Immunofluorescence and Flow Cytometry Analysis: Correlation of Leu–2, Leu–3, Leu–7, Leu–8, and Leu–11 Clee Surface Antigen Expression," *Journal of Immunology*, 132(1):151–156 (1984).

Lam et al., "A new type of synthetic peptide library for identifying ligand–binding activity," *Nature*, 354:82–84 (1991).

Laskey et al., "Messenger RNA prevalence in sea urchin embryos measured with cloned cDNAs," *PNAS*, 77(9):5317–5321 (1980).

Lee et al., "synthesis of a Polymer Surface Containing Convalently Attached Triethoxysilane Functionality: Adhesion to Glass," *Macromolecules*, 21:3353–3356 (1988).

Lehrach et al., "Labelling oligonucleotides to high specific activity (I)," *Nuc. Acids Res.*, 17(12)4605–4610 (89).

Lehrach et al., "Phage Vectors—EMBL Series," *Meth. Enzymology*, 153:103–115 (1987).

Lehrach et al., "Hybridization Fingerprinting in Genome Mapping and Sequencing," *Genome Analysis vol. 1: Genetic and Physical Mapping*, Cold Spring Harbor Laboratory Press, pp. 39–81 (1990).

Levy, M.F., "Preparing Additive Printed Circuits," *IBM Tech. Discl. Bull.*, 9(11):1473 (1967).

Lewin, Benjamin, eds., *Genes*, third edition, John Wiley & Sons, cover page, preface and table of contents, (1987).

Lichter et al., "High–Resolution Mapping of Human Chromosome 11 by in Situ hybridization with Cosmid Clones," *Science*, 247:64–69 (1990).

Lichter et al., "Fluorescence in situ hybridization with Alu and L1 polymerase chain reaction probes for rapid characterization of human chromosomes in hybrid cell lines," *PNAS*, 87:6634–6638 (1990).

Lichter et al., "Rapid detection of human chromosome 21 aberrations by in situ hybridization," *PNAS*, 85:9664–9668 (1988).

Lichter et al., "Is non–isotopic in situ hybridization finally coming of age," *Nature*, 345:93–94 (1990).

Lieberman et al., "A Light source Smaller Than the Optical Wavelength," *Science*, 247:59–61 (1990).

Lipshutz et al., "Using Oligonucleotide Probe Arrays To Access Genetic Diversity," *BioTech.*, 19(3):442–7 (1995).

Little, P., "Clone maps made simple," *Nature*, 346:611–612 (1990).

Liu et al., "Sequential Injection Analysis in Capillary Format with an Electroosmotic Pump," *Talanta*, 41(11):1903–1910 (1994).

Lockhart et al., "Expression monitoring by hybridization to high–density oligonucleotide arrays," *Nat. Biotech.*14:1675–1680 (1996).

Logue et al., "General Approaches to Mask Design for Binary Optics," SPIE, 1052:19–24 (1989).

Loken et al., "three–color Immunofluorescence Analysis of Leu Antigens on Human Peripheral Blood Using Two Lasers on a Fluorescence–Activated Cell Sorter," *Cymoetry*, 5:151–158 (1984).

Love et al., "Screening of λ Library for Differentially Expressed Genes Using in Vitro Transcripts," *Anal. Biochem.*, 150:429–441 (1985).

Lowe, C.R., "Biosensors," *Trends in Biotech.*, 2:59–65 (1984).

Lowe, C.R., "An Introduction to the Concepts and Technology of Biosensors," *Biosensors*, 1:3–16 (1985).

Lowe, C. R., Biotechnology and Crop Improvement and Protection, BCPC Publications, pp. 131–138 (1986).

Lowe et al., "Solid–Phase Optoelectronic Biosensors," *Methods in Enzymology*, 137:338–347 (1988).

Lowe, C.R., "Biosensors," *Phil. Tran. R. Soc. Lond.*, 324:487–496 (1989).

Lu et al., "Differential screening of murine ascites cDNA libraries by means of in vitro transcripts of cell–cycle–phase–specific cDNA and digital image processing," *Gene*, 86:185–192 (1990).

Luo, J. et al., "Improving the fidelity of *Thermus thermophilus* DNA ligase," *Nuc.Acids Res.*, 24(14):3071–3078 (1996).

Lysov et al., "A new method for determining the DNA nucleotide sequence by hybridization with oligonucleotides," *Doklady Biochem.*, 303(1–6):436–438 (1989).

Lysov et al., "DNA Sequencing by Oligonucleotide Hybridization," First International Conference on Electrophoresis, Supercomputingg and the Human Genome, 4/10–13/90 p. 157.

MacDonald et al., "A Rapid ELISA for Measuring Insulin in a Large Number of Research Samples," *Metabolism*, 38(5):450–452 (1989).

Mairanovsky, V.G., "Electro–Deprotection–Electrochemical Removal of Protecting Groups**," *Agnew. Chem. Int. Ed. Engl.*, 15(5):281–292 (1976).

Manz et al., "Miniaturized Total Chemical Analysis Systems: a Novel Concept for Chemical Sensing," *Sensors and Actuators*, B1:244–248 (1990).

Manz et al., "Micromachining of monocrystalline silicon and glass for chemical analysis systems, A look into next century's technology or just a fashionable craze?," *Trends in Analytical Chem.*, 10(5):144–149 (1991).

Manz et al., "Planar chips technology for minaturization and integration of separation techniques into monitoring systems, Capillary electrophoresis on a chip," *J. Chromatography*, 593:253–258 (1992).

Manz et al., "Planar Chips Technology for Miniaturization of Separation Systems: A Developing Perspective in Chemical Monitoring," chapter 1, 1–64 (1993).

Manz et al., "Electroosmotic pumping and electrophoretic separations for minaturized chemical analysis systems," *J. Micromech. Microeng.*, 4:257–265 (1994).

Masiakowski et al., "Cloning of cDNA sequences of hormone–regulated genes from the MCF–7 human breast cancer cell line," *Nuc. Acids Res.*, 10(24):7895–7903 (1982).

Matsumoto et al., "Preliminary Investigation of Micropumping Based on Electrical Control of Interfacial Tension," *IEEE*, pp. 105–110 (1990).

Matsuzawa et al., "Containment and growth of neuroblastoma cells on chemically patterned substrates," *J. Neurosci. Meth.*, 50:253–260 (1993).

Matthes et al., "Simultaneous rapid chemical synthesis of over one hundred oligonucleotides on a microscale," *EMBO J.*, 3(4):801–805 (1984).

McCray et al., "Properties and Uses of Photoreactive Caged Compounds," *Ann. Rev. Biophys. Biophys. Chem.*, 18:239–270 (1989).

McGall et al., "The Efficiency of Light–Directed Synthesis of DNA Arrays on Glass Substrates," *J. American Chem. Soc.*, 119(22):5081–5090 (1997).

McGillis, VLSI Technology, Sze, eds., Chapter 7, "Lithography," pp. 267–301 (1983).

McMurray, J.S., "Solid Phase Synthesis of a Cyclic Peptide Using Fmoc Chemistry," *Tetrahedron Letters*, 32(52):7679–7682 (1991).

Meinkoth et al., "Review: Hybridization of Nucleic Acids Immobilized on solid Supports," *Analytical Biochem.*, 138:267–284 (1984).

Melcher et al., "Traveling–Wave Bulk Electroconvection Induced across a Temperature Gradient," *Physics of Fluids*, 10(6):1178–1185 (1967).

Merrifield, R.B., "Solid Phase peptide Synthesis. I. The Synthesis of a Tetrapeptide," *J.Am.Chem.Soc.*, 85:2149–2154 (1963).

Michiels et al., "Molecular approaches to genome analysis: a strategy for the construction of ordered overlapping clone libraries," *CABIOS*, 3(3):203–10 (1987).

Mirzabekov, A.D., "DNA sequencing by hybridization—a megasequencing method and a diagnostic tool?," *TIBTECH*, 12:27–32 (1994).

Miyada et al., "Oligonucleotide Hybridization Techniques," *Meth. Enzymology*, 154:94–107 (1987).

Monaco et al., "Human Genome Linking with Cosmids and Yeast Artificial Chromosomes", abstract from CSHS, p. 50, (1989).

Morita et al., "Direct pattern fabrication on silicone resin by vapor phase electron beam polymerization," *J.Vac.Sci.Technol.*, B1(4):1171–1173 (1983).

Morrison et al., "Solution–Phase Detection of Polynucleotides Using Interacting Fluorescent Labels and Competitive Hybridization," *Anal. Biochem.*, 183:231–244 (1989).

Munegumi et al., "thermal Synthesis of Polypeptides from N–Boc–Amino Acid (Aspartic Acid, β–Aminoglutaric Acid) Anhydrides," *Chem. Letters*, pp. 1643–1646 (1988).

Mutter et al., "Impact of Conformation on the Synthetic Strategies for Peptide Sequences," pp. 217–228 from Chemistry of Peptides and Proteins, vol. 1, Proceedings of the Third USSR–FRG Symp., in USSR (1982).

Nakamori et al., "A Simple and Useful Method for Simultaneous Screening of Elevated Levels of Expression of a Variety of Oncogenes in Malignant Cells," *Jpn. J. Cancer Res.*, 79:1311–1317 (1988).

Nederlof et al., "Multiple Fluorescence In Situ Hybridization," *Cytometry*, 11:126–131 (1990).

Nederlof et al., "Three–Color Fluorescence In Situ Hybridization for the Simultaneous Detection of Multiple Nucleic Acids Sequences," *Cytometry*, 10:20–27 (1989).

Nizetic et al., "An improved bacterial colony lysis procedure enables direct DNA hybridisation using short (10, 11 bases) oligonucleotides to cosmids," *Nuc. Acids Res.*, 19(1):182 (1990).

Nizetic et al., "Construction, arraying, and high–density screening of large insert libraries of human chromosomes X and 21: their potential use as reference libraries," *PNAS*, 88:3233–3237 (1991).

Nyborg, W., "Acoustic Streaming," chapter 11 pp. 265–329 from Physical Acoustics, Principles and Methods, Mason, eds., vol. II, part B, Academic Press, New York and London (1965).

Ocvirk et al., "High Performance Liquid Chromatography Partially Integrated onto a Silicon Chip," *Analyt. Meth. Instrumentation*, 2(2):74–82 (1995).

Ohtsuka et al., "Studies on transfer ribonucleic acids and related compounds. IX Ribonucleic oligonucleotide synthesis using a photosensitive 0–nitrobenzyl protection at the 2'–hydroxl group," *Nuc.Acids.Res.*, 1(10):1351–1357 (1974).

Olefirowicz et al., "Capillary Electrophoresis for Sampling Single Nerve Cells," *Chimia*, 45(4):106–108 (1991).

Olson et al., "Random–clone strategy for genomic restriction mapping in yeast," *PNAS*, 83:7826–7830 (1986).

Patchornik et al., "Photosensitive Protecting Groups," *J.Am. Chem.Soc.*, 92(21):6333–6335 (1970).

Patent Abstracts of Japan from EPO, Abst. 13:557, JP 1–233 447 (1989).

Pease et al., "Light–generated oligonucleotide arrays for rapid DNA sequence analysis," *PNAS*, 91:5022–26 (1994).

Pevzner, P.A., "DNA Physical Mapping and Alternating Eulerian Cycles in Colored Grapes," *Algorithmica*, 13(1–2):77–105 (1995).

Pevzner et al., "Multiple Filtration and Approximate Pattern Matching," *Algorithmica*, 13(1–2):135–154 (1995).

Pevzner et al., "Generalized Sequence Alignment and Duality," *Adv. Applied Math.*, 14:139–171 (1993).

Pevzner, P.A., "1–Tuple DNA Sequencing: Computer Analysis," *J. Biomol. Struct. Dynam.*, 7(1):63–69 (1989).

Pfahler et al., "Liquid Transport in Micron and Submicron Channels," *Sensors and Actuators*, A21–A23:431–4 (90).

Pfeifer et al., "Genomic Sequencing and Methylation Analysis by Ligation Mediated PCR," *Science*, 246:810–813 (1989).

Pidgeon et al., "Immobilized Artificial Membrane Chromatography: Supports Composed of Membrane Lipids," *Anal. Biochem.*, 176:36–47 (89).

Pillai, V.N., "Photoremovable Protecting Groups in Organic Synthesis," *Synthesis*, pp. 1–26 (1980).

Pillai et al., "3–Nitro–4–Aminomethylbenzoylderivate von Polyethylenglykolen: Eine neue Klasse von Photosensitiven loslichen Polymeren Tragern zur Synthese von C–terminalen Peptidamiden," *Tetrah. ltr.*, # 36 p. 3409–3412 (1979).

Pillai et al., "Synthesis Hydrophilic Polymers, Biomedical and Chemical Applications," *Naturwissenschaften*, 68:558–566 (1981).

Pirrung et al., "Proofing of Photolithographic DNA Synthesis with 3',5'–Dimethoxybenzoinyloxycarbonyl–Protected Deoxynucleoside Phosphoramidites," *J. Org. Chem.*, 63(2):241–246 (1998).

Pirrung et al., "Comparison of Methods for Photochemical Phosphoramidite–Based DNA Synthesis," *J. Org. Chem.*, 60:6270–6276 (1995).

Ploax et al., "Cyclization of peptides on a solid support," *Int. J. Peptide Protein Research*, 29:162–169 (1987).

Polsky–Cynkin et al., "Use of DNA Immobilized on Plastic and Agarose Supports to Detect DNA by Sandwich Hybridization," *Clin. Chem.*, 31(9):1428–1443 (1985).

Poustka et al., "Molecular Approaches to Mammalian Genetics," Cold Spring Harbor Symposia on Quantitive Biology, 51:131–139 (1986).

Purushothaman et al., "Synthesis of 4,5–diarylimidazoline–2 thiones and their photoconversion to bis(4, 5–diarylimidazol–2–yl) sulphides," *Ind. J. Chem.*, 29B:18–21 (1990).

Quesada et al., "High–Sensitivity DNA Detection with a Laser–Exited Confocal Fluorescence Gel Scanner," *Biotechniques*, 10:616 (1991).

Reichmanis et al., *J. Polymer Sci. Polymer Chem. Edition*, 23:1–8 (1985).

Renz et al., "A colorimetric method for DNA hybridization," *Nuc. Acids Res.*, 12(8):3435–3445 (1984).

Richter et al., "An Electrohydrodynamic Micropump," *IEEE*, pp. 99–104 (1990).

Richter et al., "Electrohydrodynamic Pumping and Flow Measurement," *IEEE*, pp. 271–276 (1991).

Richter et al., "A Micromachined electrohydrodynamic (EHD) pump," *Sensors and Actuators*, A29:159–168 (91).

Robertson et al., "A General and Efficient Route for Chemical Aminoacylation of Transfer RNAs," *J. Am. Chem. Soc.*, 113:2722–2729 (1991).

Rodda et al., "The Antibody Response to Myoglobin–I. Systematic Synthesis of Myglobin Peptides Reveals Location and Substructure of Species–Dependent Continuous Antigenic Determinants," *Mol. Immunol.*, 23(6):603–610 (1986).

Rodgers, R.P., "Data Processing of Immunoassay Results," Manual of Clin. Lab. Immunol., 3rd ed., ch. 15, pp. 82–87 (1986).

Rose, D.J., "Free–solution reactor for post–column fluorescence detection in capillary zone electrophoresis," *J. Chromatography*, 540:343–353 (1991).

Rovero et al., "Synthesis of Cyclic Peptides on solid Support," *Tetrahed. Letters*, 32(23):2639–2642 (1991).

Sambrook, Molecular Cloning—A Laboratory Manual, publ. in 1989 (not included).

Saiki et al., "Genetic analysis of amplified DNA with immobilized sequence–specific oligonucleotide probes," *PNAS*, 86:6230–6234 (1989).

Saiki et al., "Analysis of enzymatically amplified β–globin and HLA–DOα DNA with Allele–specific oligonucleotide probes," *Nature*, 324:163–166 (1986).

Schafer et al., "DNA fingerprinting using non–radioactive oligonucleotide probes specific for simple repeats," *Nuc. Acids Res.*, 16(19):9344 (1988).

Scharf et al., "HLA class II allelic variation and susceptibility to pemphigus vulgaris," *PNAS*, 85(10):3504–3508 (1988).

Schena et al., "Paralle human genome analysis: Microarray–based expression monitoring of 1000 genes," *PNAS*, 93:10614–10619 (1996).

Schuup et al., "Mechanistic Studies of the Photorarrangement of o–Nitrobenzyl Esters," *J. Photochem.*, 36:85–97 (1987).

Seed, B., "Diazotizable arylamine cellulose papers for the coupling and hybridization of nucleic acids," *Nuc. Acids Res.*, 10(5):1799–1810 (1982).

Seiler et al., "Planar Glass Chips for Capillary Electrophoresis: Repetitive Sample Injection, Quantitation, and Separation Efficency," *Anal. Chem.*, 65:1481–1488 (1993).

Seller et al., "Electroosmotic Pumping and Valveless Control of Fluid Flow with a Manifold of Capillaries on a Glass Chip," Anal. Chem., 66:3485–3491 (1994).

Semmelhack et al., "Selective Removal of Protecting Groups Using Controlled Potential Electrolysis," *J. Am. Chem. Society*, 94(14):5139–5140 (1972).

Sheldon et al., "Matrix DNA Hybridization," *Clinical Chemistry*, 39(4):718–719 (1993).

Shin et al., "Dehydrooligonpeptides. XI. Facile Synthesis of Various Kinds of Dyhydrodi–and tripeptides, and Dehydroenkephalins Containing Tyr Residue by Using N–Carboxydehydrotyrosine Anhydride," *Bull. Chem. Soc. Jpn.*, 62:1127–1135 (1989).

Sim et al., "Use of a cDNA Library for Studies on Evolution and Developmental Expression of the Chorion Multigene Families," *Cell*, 18:1303–1316 (1979).

Smith et al., "A Novel Method for Delineating Antigenic Determinants: Peptide Synthesis and Radioimmunoassay Using the Same Solid Support," *Immunochemistry*, 14:565–568 (1977).

Sofia, M.J., "Carbohydrate–based combinatorial libraries," *Molecular Diversity*, 3:75–94 (1998).

Southern et al., "Report on the Sequencing by Hybridization Workshop," *Genomics*, 13:1378–1383 (1992).

Southern et al., "Oligonucleotide hybridisations on glass supports: a novel linker for oligonucleotide synthesis and hybridization properties of oligonucleotides synthesized in situ," *Nuc. Acids Res.*, 20(7):1679–1684 (1992).

Southern et al., "Analyzing and Comparing Nucleic Acid Sequences by Hybridization to Arrays of Oligonucleotides: Evaluation Using Experimental Models," Genomics, 13:1008–10017 (1992).

Southern, E.M., "Detection of Specific Sequences Among DNA Fragments Separted by Gel Electrophoresis," *J. Mol. Biol.*, 98:503–517 (1975).

Stemme et al., "A valveless diffuser/nozzle–based fluid pump," *Sensors and Actuators*, A39:159–167 (1993).

Stryer, L., "DNA Probes and Genes Can be Synthesized by Automated Solid–Phase Methods," from *Biochemistry*, Third Edition, published by W.H. Freeman & Co., (1988).

Stuber et al., "Synthesis and photolytic cleavage of bovine insulin B22–30 on a nitrobenzoylglycyl–poly (ethylene glycol) support," *Int. J. Peptide Protein Res.*, 22(3):277–283 (1984).

Sundberg et al., "Spatially–Addressable Immobilization of Macromolecules on Solid Supports," *J. Am. Chem. Soc.*, 117(49):12050–12057 (1995).

Swedberg, S.A., "Use of non–ionic and zwitterionic surfactants to enhance selectivity in high–performance capillary electrophoresis, An apparent micellar electrokinetic capillary chromatography mechanism," *J. Chromatography*, 503:449–452 (1990).

Thomas, P.S., "Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose," *PNAS*, 77(9):5201–5205 (1980).

Titus et al., "Texas Red, a Hydrophilic, red–emitting fluorophore for use with fluorescein in dual parameter plow microfluorometric and fluorescence microscopic studies," *J. Immunol. Meth.*, 50:193–204 (1982).

Tkachuk et al., "Detection of bcr–abl Fusion in chronic Myelogeneous Leukemia by in situ Hybridization," *Science*, 250:559–562 (90).

Trzeciak et al., "Synthesis of 'Head–to–Tail' Cyclized Peptides on Solid Support by FMOC Chemistry," *Tetrahed. Letters*, 33(32):4557–4560 (1992).

Tsien et al., "Control of Cytoplasmic Calcium with Photolabile Tetracarboxylate 2–Nitrobenzhydrol Chelators," *Biophys. J.*, 50:843–853 (1986).

Tsutsumi et al., "Expression of L– and M– Type Pyruvate Kinase in Human Tissues," *Genomics*, 2:86–89 (1988).

Turchinskii et al., "Multiple Hybridization in Genome Analysis, Reaction of Diamines nd Bisulfate with Cytosine for Introduction of Nonradioactive labels Into DNA," *Molecular Biology*, 22:1229–1235 (1988).

Turner et al., "Photochemical Activation of Acylated α–Thrombin," *J. Am. Chem. Soc.*, 109:1274–1275 (1987).

Urdea et al., "A novel method for the rapid detection of specific nucleotide sequences in crude biological samples without blotting or radioactivity; application to the analysis of hepatitis B virus in human serum," *Gene*, 61:253–264 (1987).

Urdea et al.,"A comparison of non–radioisotopic hybridization assay methods using fluorescent, chemiluminescent and enzyme labeled synthetic oligodeoxyribonucleotide probes," *Nuc. Acids. Res.*, 16(11):4937–4956 (1988).

Van der Voort et al., "Design and Use of a Computer Controlled Confocal Microscope for Biological Applications," *Scanning*, 7(2):66–78 (1985).

Van Hijfte et al., "Intramolecular 1,3–Diyl Trapping Reactions. A Formal Total Synthesis of –Coriolin," J. Organic Chemistry, 50:3942–3944 (1985).

Veldkamp, W.B., "Binary optics: the optics technology of the 1990s," CLEO 90, vol. 7, paper # CMG6 (1990).

Verlaan–de Vries et al., "A dot–blot screening procedure for mutated ras oncogenes using synthetic oligodeoxynucleotides," *Gene*, 50:313–320 (1986).

Verpoorte et al., "Three–dimensional micro flow manifolds for miniaturized chemical analysis systems," *J. Micromech. Microeng.*, 4:246–256 (1994).

Volkmuth et al., "DNA electrophoresis in microlithographic arrays," *Nature*, 358:600–602 (1992).

Voss et al., "The immobilization of oligonucleotides and their hybridization properties," *Biochem. Soc. Transact.*, 16:216–217 (1988).

Wada, A., *International Workshop on Automatic and High Speed DNA Base Sequencing*, Hayashibara Forum 1987 at Hayashibara Biochemical Laboratories, Okayama, Japan, Jul. 7–9, 1987.

Walker et al., "Photolabile Protecting Groups for an Acetylcholine Receptor Ligand. Synthesis and Photochemistry of a New Class of o–Nitrobenzyl Derivatives and their Effects on Receptor Function," *Biochemistry*, 25:1799–1805 (1986).

Wallace et al., "The use of synthetic oligonucleotides as hybridization probes. II. Hybridization of oligonucleotides of mixed sequence to rabbit β–globoin DNA," *Nuc. Acids Res.*, 9(4):879 (1981).

Wallace et al., "Hybridization of synthetic oligodeoxyribonucleotides to ΦX 174 DNA: the effect of single base pair mismatch," *Nuc. Acids Res.*, 11(6):3543–3557 (1979).

Washizu et al., "Handling Biological Cells Using a Fluid Integrated Circuit," *IEEE Transactions Industry Applications*, 26(2):352–358 (1990).

Wiedmann, M. et al., "Ligase Chain Reaction (LCR)—Overview and Applications," *PCR Meth. Appl.*, 3(4):S51–S64 (1994).

Werner et al., "Size–Dependent Separation of Proteins Denatured in SDS by Capillary Electrophoresis Using a Replaceable Sieving Matrix," *Anal. Biochem.*, 212:253–258 (1993).

White et al., "An Evaluation of Confocal Versus Conventional Imaging of Biological Structures by Fluorescence Light Microscopy," *J. Cell Biol.*, 105(1):41–48 (1987).

Widacki et al., "Biochemical Differences in Qa–2 Antigens Expressed by Qa–2+,6+ and Qa–2a+,6– Strains. Evidence for Differential Expression of the Q7 and Q9 Genes," *Mol. Immunology*, 27(6):559–570 (1990).

Wilcox et al., "Synthesis of Photolabile 'Precursors' of Amino Acid Neurotransmitters," *J. Org. Chem.*, 55:1585–1589 (1990).

Wilding et al., "PCR in a Silicon Microstructure," *Clin. Chem.*, 40(9):1815–1818 (1994).

Wilding et al., "Manipulation and Flow of Biological Fluids in Straight Channels Micromachined in Silicon," *Clin. Chem.*, 40(1):43–47 (1994).

Wittman–Liebold, eds., Methods in Protein Sequence Analysis, from Proceedings of 7th Int'l Conf., Berlin, Germany, 7/3–8/88, table of contents, pp. xi–xx* (1989).

Wood et al., "Base composition–independent hybridization in tetramethylammonium chloride: A method for oligonucleotide screening of highly complex gene libraries," *PNAS*, 82:1585–1588 (1985).

Woolley et al., "Ultra–high–speed DNA fragment separations using microfabricated capillary array electrophoresis chips," *PNAS*, 91:11348–11352 (1994).

Wu et al., "Synthesis and Properties of Adenosine–5'–triphosphoro–γ–5–(5–sulfonic acid)naphthyl Ethylamidate: A Fluorescent Nucleotide Substrate for DNA–Dependent RNA Polymerase from *Escherichia coli*," *Arch. Biochem. Biophys.*, 246(2):564–571 (1986).

Wu et al., "Laboratory Methods, Direct Analysis of Single Nucleotide Variation in Human DNA and RNA Using In Situ Dot Hybridization," *DNA*, 8(2):135–142 (1989).

Yamamoto et al., "Features and applications of the laser scanning microscope," *J. Mod. Optics*, 37(11):1691–1701 (1990).

Yarbrough et al., "Synthesis and Properties of Fluorescent Nucleotide Substrates for DNA–dependent RNA Polymerases," *J. Biol. Chem.*, 254(23):12069–12073 (1979).

Yosomiya et al., "Performance, Glass fiber Having Isocyanate Group on the Surface. Preparation and Reaction with Amino Acid," *Polymer Bulletin*, 12:41–48 (1984).

Young, W.S.,"Simultaneous Use of Digoxigenin– and Radiolabeled Oligodeoxyribonucleotide Probes for Hybridization Histochemistry," *Neuropeptides*, 13:271–275 (1989).

Yue et al., "Miniature Field–Flow Fractionation System for Analysis of Blood Cells," *Clin. Chem.*, 40(9):1810–1814 (1994).

Zehavi et al., "Light–Sensitive Glycosides. I. 6–Nitroveratryl β–D–Glucopyranoside and 2–Nitrobenzyl β–D–Glucopyranoside," *J. Org. Chem.*, 37(14):2281–2285 (1972).

Zengerle et al., "Transient measurements on miniaturized diaphragm pumps in microfluid systems," *Sensors and Actuators*, A46–47:557–561 (1995).

Zischler et al., "Non–radioacive oligonucleotide fingerprinting in the gel," *Nuc. Acids Res.*, 19(11)4411 (1989).

Zischler et al., "Digoxigenated oligonucleotide probes specific for simple repeats in DAN fingerprinting and hybridization in situ," *Hum. Genet.*, 82:227–233 (1989).

Hodgson et al, Nucl. Acids Res., 15(15):6295 (1987).

Khrapko et al, DNA Seq. Map, 1:375–388 (1991).

Lander et al, Genomics, 2:231–239 (1988).

Little, Nature, 346:611–612 (1990).

Lysov et al, Dokl. Akad. Nauk. SSSR, 303:1508–1511 (1988).

Olson et al, Proc. Natl. Acad. Sci. USA, 83:7826–7830 (Oct. 1986).

Pevzner, Algorithmica, 13(1–2):77–105 (1995).

Pevzner et al, Algorithmica, 13(1–2):135–154 (1995).

Pfeifer et al, Science, 246:810–813 (Nov. 10, 19889).

Seed, Nucl. Acids Res., 10(5):1799–1810 (1982).

Wood et al, Proc. Natl. Acad. Sci. USA, 82:1585–1588 (1985).

Feinberg et al, Anal. Biochem., 137:266–267 (1984).

Pevzner et al, Adv. Applied Math, 14:139–171 (1993).

Schena et al, Proc. Natl. Acad. Sci. USA, 93:10614–10619 (Oct. 1996).

Miller et al. "Detection of bacteria by hybridization of rRNA with DNA–latex and immunodetection of hybrids" J Clin Microbiol 1988, 26:1271–1276.

Brenner et al., "In vitro cloning of complex mixtures of DNA on microbeads: Physical separation of differentially expressed cDNAs", PNAS, vol. 97, No. 4, Feb. 15, 2000, pp. 1665–1670.

Brenner et al., "Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays", Nature Biotechnology, vol. 18, Jun. 2000, pp. 630–634.

Tyagi, "Taking a census of mRNA populations with microbeads", Nature Biotechnology, vol. 18, Jun. 2000, pp. 597 and 598.

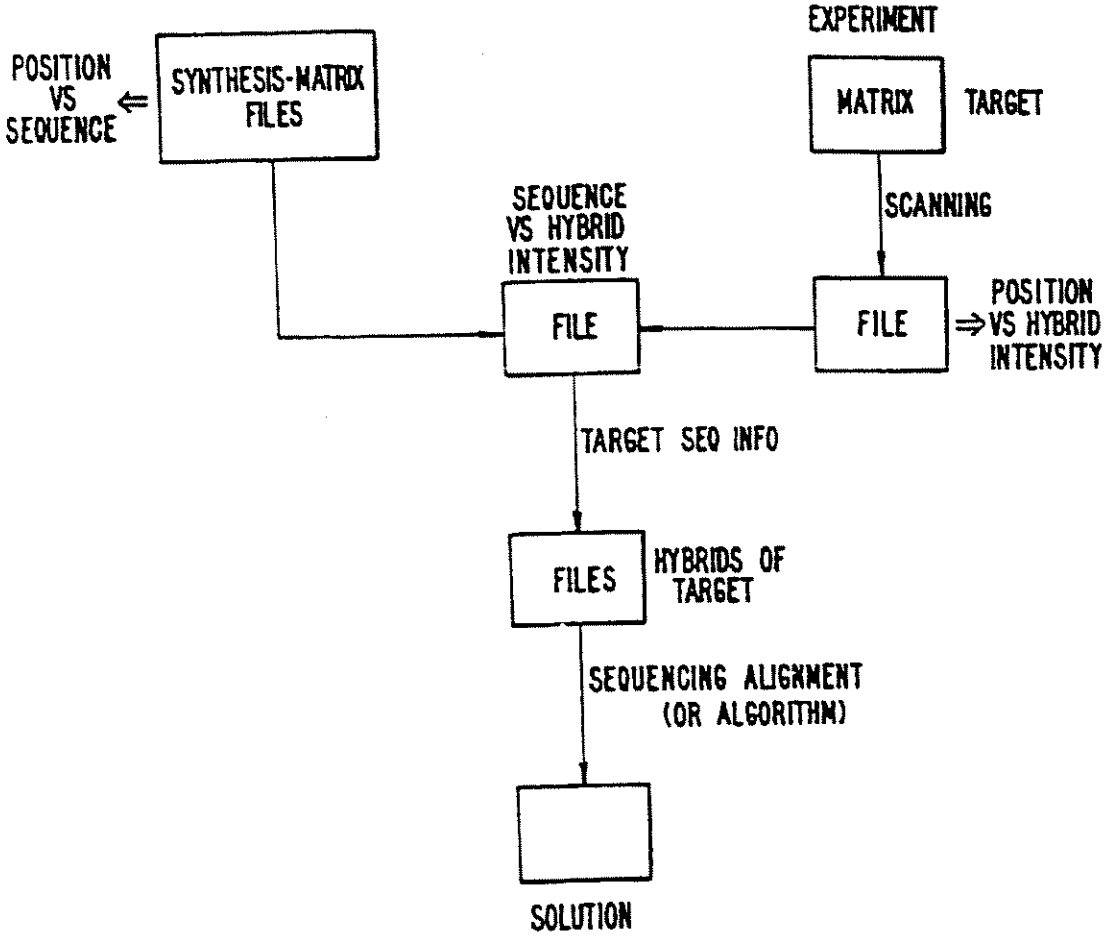
EXPERIMENT
POSITION VS SEQUENCE
SYNTHESIS-MATRIX FILES
MATRIX
TARGET
SCANNING
SEQUENCE VS HYBRID INTENSITY
FILE
FILE
POSITION VS HYBRID INTENSITY
TARGET SEQ INFO
FILES
HYBRIDS OF TARGET
SEQUENCING ALIGNMENT (OR ALGORITHM)
SOLUTION

FIG. 1

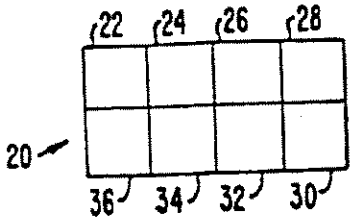

FIG. 2A

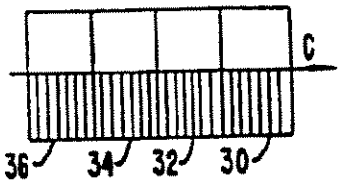

FIG. 2B

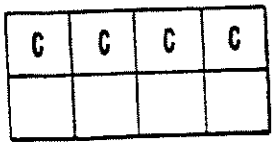

FIG. 2C

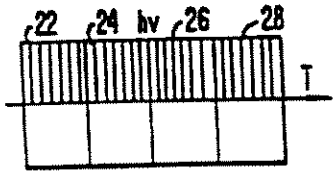

FIG. 2D

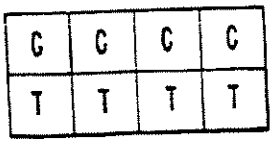

FIG. 2E

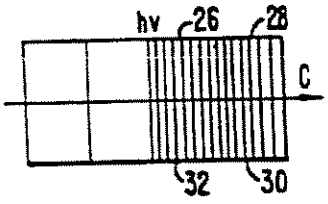

FIG. 2F

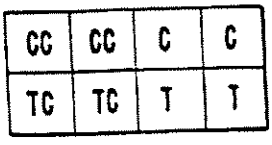

FIG. 2G

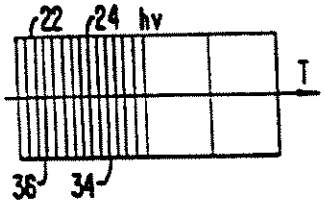

FIG. 2H

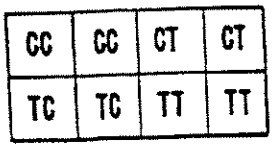

FIG. 2I

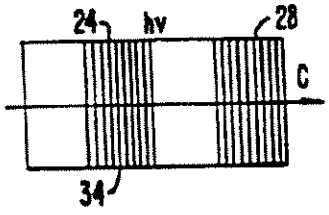

FIG. 2K

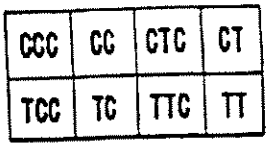

FIG. 2K

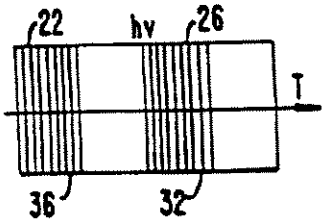

FIG. 2L

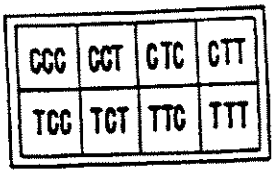

FIG. 2M

PRODUCTS FOR DETECTING NUCLEIC ACIDS

CROSS REFERENCE TO RELATED APPLICATIONS

This is a continuation of application Ser. No. 09/362,089, filed Jul. 28, 1998, pending; which is a divisional of application Ser. No. 09/056,927, filed Apr. 8, 1998, now U.S. Pat. No. 6,197,506; which is a continuation of application Ser. No. 08/670,118, filed Jun. 25, 1996, now U.S. Pat. No. 5,800,992; which is a divisional of application Ser. No. 08/168,904, filed Dec. 15, 1993, now abandoned; which is a continuation of application Ser. No. 07/624,114, filed Dec. 6, 1990, now abandoned; each of which is hereby incorporated by reference.

Additional commonly assigned application No. 07/492,462, filed Mar. 7, 1990, now U.S. Pat. No. 5,143,854; application No. 07/362,901, filed Jun. 7, 1989, now abandoned; application Ser. Nos. 07/624,120 and 07/626,730, both of which were filed on Dec. 6, 1990; application Ser. No. 07/435,316, filed Nov. 13, 1989, now abandoned; and U.S. Pat. No. 5,252,743 are also hereby incorporated herein by reference.

BACKGROUND OF THE INVENTION

The present invention relates to the sequencing, fingerprinting, and mapping of polymers, particularly biological polymers. The inventions may be applied, for example, in the sequencing, fingerprinting, or mapping of nucleic acids, polypeptides, oligosaccharides, and synthetic polymers.

The relationship between structure and function of macromolecules is of fundamental importance in the understanding of biological systems. These relationships are important to understanding, for example, the functions of enzymes, structural proteins, and signalling proteins, ways in which cells communicate with each other, as well as mechanisms of cellular control and metabolic feedback.

Genetic information is critical in continuation of life processes. Life is substantially informationally based and its genetic content controls the growth and reproduction of the organism and its complements. Polypeptides, which are critical features of all living systems, are encoded by the genetic material of the cell. In particular, the properties of enzymes, functional proteins, and structural proteins are determined by the sequence of amino acids which make them up. As structure and function are integrally related, many biological functions may be explained by elucidating the underlying structural features which provide those functions. For this reason, it has become very important to determine the genetic sequences of nucleotides which encode the enzymes, structural proteins, and other effectors of biological functions. In addition to segments of nucleotides which encode polypeptides, there are many nucleotide sequences which are involved in control and regulation of gene expression.

The human genome project is directed toward determining the complete sequence of the genome of the human organism. Although such a sequence would not correspond to the sequence of any specific individual, it would provide significant information as to the general organization and specific sequences contained within segments from particular individuals. It would also provide mapping information which is very useful for further detailed studies. However, the need for highly rapid, accurate, and inexpensive sequencing technology is nowhere more apparent than in a demanding sequencing project such as this. To complete the sequencing of a human genome would require the determination of approximately 3×10^9, or 3 billion base pairs.

The procedures typically used today for sequencing include the Sanger dideoxy method, see, e.g., Sanger et al. (1977) *Proc. Natl. Acad. Sci. USA*, 74:5463–5467, or the Maxam and Gilbert method, see, e.g., Maxam et al., (1980) *Methods in Enzymology*, 65:499–559. The Sanger method utilizes enzymatic elongation procedures with chain terminating nucleotides. The Maxam and Gilbert method uses chemical reactions exhibiting specificity of reaction to generate nucleotide specific cleavages. Both methods require a practitioner to perform a large number of complex manual manipulations. These manipulations usually require isolating homogeneous DNA fragments, elaborate and tedious preparing of samples, preparing a separating gel, applying samples to the gel, electrophoresing the samples into this gel, working up the finished gel, and analyzing the results of the procedure.

Thus, a less expensive, highly reliable, and labor efficient means for sequencing biological macromolecules is needed. A substantial reduction in cost and increase in speed of nucleotide sequencing would be very much welcomed. In particular, an automated system would improve the reproducibility and accuracy of procedures. The present invention satisfies these and other needs.

SUMMARY OF THE INVENTION

The present invention provides improved methods useful for de novo sequencing of an unknown polymer sequence, for verification of known sequences, for fingerprinting polymers, and for mapping homologous segments within a sequence. By reducing the number of manual manipulations required and automating most of the steps, the speed, accuracy, and reliability of these procedures are greatly enhanced.

The production of a substrate having a matrix of positionally defined regions with attached reagents exhibiting known recognition specificity can be used for the sequence analysis of a polymer. Although most directly applicable to sequencing, the present invention is also applicable to fingerprinting, mapping, and general screening of specific interactions. The VLSIPS™ Technology (Very Large Scale Immobilized Polymer Synthesis) substrates will be applied to evaluating other polymers, e.g., carbohydrates, polypeptides, hydrocarbon synthetic polymers, and the like. For these non-polynucleotides, the sequence specific reagents will usually be antibodies specific for a particular subunit sequence.

According to one aspect of the masking technique, the invention provides an ordered method for forming a plurality of polymer sequences by sequential addition of reagents comprising the step of serially protecting and deprotecting portions of the plurality of polymer sequences for addition of other portions of the polymer sequences using a binary synthesis strategy.

The present invention also provides a means to automate sequencing manipulations. The automation of the substrate production method and of the scan and analysis steps minimizes the need for human intervention. This simplifies the tasks and promotes reproducibility.

The present invention provides a composition comprising a plurality of positionally distinguishable sequence specific reagents attached to a solid substrate, which reagents are capable of specifically binding to a predetermined subunit sequence of a preselected multi-subunit length having at

least three subunits, said reagents representing substantially all possible sequences of said preselected length. In some embodiments, the subunit sequence is a polynucleotide or a polypeptide, in others the preselected multi-subunit length is five subunits and the subunit sequence is a polynucleotide sequence. In other embodiments, the specific reagent is an oligonucleotide of at least about five nucleotides. Alternatively, the specific reagent is a monoclonal antibody. Usually the specific reagents are all attached to a single solid substrate, and the reagents comprise about 3000 different sequences. In other embodiments, the reagents represents at least about 25% of the possible subsequences of said preselected length. Usually, the reagents are localized in regions of the substrate having a density of at least 25 regions per square centimeter, and often the substrate has a surface area of less than about 4 square centimeters.

The present invention also provides methods for analyzing a sequence of a polynucleotide or a polypeptide, said method comprising the step of:

a) exposing said polynucleotide or polypeptide to a composition as described.

It also provides useful methods for identifying or comparing a target sequence with a reference, said method comprising the step of:

a) exposing said target sequence to a composition as described;

b) determining the pattern of positions of the reagents which specifically interact with the target sequence; and

c) comparing the pattern with the pattern exhibited by the reference when exposed to the composition.

The present invention also provides methods for sequencing a segment of a polynucleotide comprising the steps of:

a) combining:

i) a substrate comprising a plurality of chemically synthesized and positionally distinguishable oligonucleotides capable of recognizing defined oligonucleotide sequences; and

ii) a target polynucleotide; thereby forming high fidelity matched duplex structures of complementary subsequences of known sequence; and

b) determining which of said reagents have specifically interacted with subsequences in said target polynucleotide.

In one embodiment, the segment is substantially the entire length of said polynucleotide.

The invention also provides methods for sequencing a polymer, said method comprising the steps of:

a) preparing a plurality of reagents which each specifically bind to a subsequence of preselected length;

b) positionally attaching each of said reagents to one or more solid phase substrates, thereby producing substrates of positionally definable sequence specific probes;

c) combining said substrates with a target polymer whose sequence is to be determined; and

d) determining which of said reagents have specifically interacted with subsequences in said target polymer.

In one embodiment, the substrates are beads. Preferably, the plurality of reagents comprise substantially all possible subsequences of said preselected length found in said target. In another embodiment, the solid phase substrate is a single substrate having attached thereto reagents recognizing substantially all possible subsequences of preselected length found in said target.

In another embodiment, the method further comprises the step of analyzing a plurality of said recognized subsequences to assemble a sequence of said target polymer. In a bead embodiment, at least some of the plurality of substrates have one subsequence specific reagent attached thereto, and the substrates are coded to indicate the sequence specificity of said reagent.

The present invention also embraces a method of using a fluorescent nucleotide to detect interactions with oligonucleotide probes of known sequence, said method comprising:

a) attaching said nucleotide to a target unknown polynucleotide sequence, and

b) exposing said target polynucleotide sequence to a collection of positionally defined oligonucleotide probes of known sequences to determine the sequences of said probes which interact with said target.

In a further refinement, an additional step is included of:

a) collating said known sequences to determine the overlaps of said known sequences to determine the sequence of said target sequence.

A method of mapping a plurality of sequences relative to one another is also provided, the method comprising:

a) preparing a substrate having a plurality of positionally attached sequence specific probes;

b) exposing each of said sequences to said substrate, thereby determining the patterns of interaction between said sequence specific probes and said sequences; and

c) determining the relative locations of said sequence specific probe interactions on said sequences to determine the overlaps and order of said sequences.

In one refinement, the sequence specific probes are oligonucleotides, applicable to where the target sequences are nucleic acid sequences.

In the nucleic acid sequencing application, the steps of the sequencing process comprise:

a) producing a matrix substrate having known positionally defined regions of known sequence specific oligonucleotide probes;

b) hybridizing a target polynucleotide to the positions on the matrix so that each of the positions which contain oligonucleotide probes complementary to a sequence on the target hybridize to the target molecule;

c) detecting which positions have bound the target, thereby determining sequences which are found on the target; and

d) analyzing the known sequences contained in the target to determine sequence overlaps and assembling the sequence of the target therefrom.

The enablement of the sequencing process by hybridization is based in large part upon the ability to synthesize a large number (e.g., to virtually saturate) of the possible overlapping sequence segments and distinguishing those probes which hybridize with fidelity from those which have mismatched bases, and to analyze a highly complex pattern of hybridization results to determine the overlap regions.

The detecting of the positions which bind the target sequence would typically be through a fluorescent label on the target. Although a fluorescent label is probably most convenient, other sorts of labels, e.g., radioactive, enzyme linked, optically detectable, or spectroscopic labels may be used. Because the oligonucleotide probes are positionally defined, the location of the hybridized duplex will directly translate to the sequences which hybridize. Thus, analysis of the positions provides a collection of subsequences found within the target sequence. These subsequences are matched

with respect to their overlaps so as to assemble an intact target sequence.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 illustrates a flow chart for sequence, fingerprint, or mapping analysis.

FIGS. 2A–2M illustrate the proper function of a VLSIPS™ Technology nucleotide synthesis.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. OVERALL DESCRIPTION

A. General

The present invention relies in part on the ability to synthesize or attach specific recognition reagents at known locations on a substrate, typically a single substrate. In particular, the present invention provides the ability to prepare a substrate having a very high density matrix pattern of positionally defined specific recognition reagents. The reagents are capable of interacting with their specific targets while attached to the substrate, e.g., solid phase interactions, and by appropriate labeling of these targets, the sites of the interactions between the target and the specific reagents may be derived. Because the reagents are positionally defined, the sites of the interactions will define the specificity of each interaction. As a result, a map of the patterns of interactions with specific reagents on the substrate is convertible into information on the specific interactions taking place, e.g., the recognized features. Where the specific reagents recognize a large number of possible features, this system allows the determination of the combination of specific interactions which exist on the target molecule. Where the number of features is sufficiently large, the identical same combination, or pattern, of features is sufficiently unlikely that a particular target molecule may often be uniquely defined by its features. In the extreme, the features may actually be the subunit sequence of the target molecule, and a given target sequence may be uniquely defined by its combination of features.

In particular, the methodology is applicable to sequencing polynucleotides. The specific sequence recognition reagents will typically be oligonucleotide probes which hybridize with specificity to subsequences found on the target sequence. A sufficiently large number of those probes allows the fingerprinting of a target polynucleotide or the relative mapping of a collection of target polynucleotides, as described in greater detail below.

In the high resolution fingerprinting provided by a saturating collection of probes which include all possible subsequences of a given size, e.g., 10-mers, collating of all the subsequences and determination of specific overlaps will be derived and the entire sequence can usually be reconstructed.

Although a polynucleotide sequence analysis is a preferred embodiment, for which the specific reagents are most easily accessible, the invention is also applicable to analysis of other polymers, including polypeptides, carbohydrates, and synthetic polymers, including α-, β-, and ω-amino acids, polyurethanes, polyesters, polycarbonates, polyureas, polyamides, polyethyleneimines, polyarylene sulfides, polysiloxanes, polyimides, polyacetates, and mixed polymers. Various optical isomers, e.g., various D- and L-forms of the monomers, may be used.

Sequence analysis will take the form of complete sequence determination, to the level of the sequence of individual subunits along the entire length of the target sequence. Sequence analysis also takes the form of sequence homology, e.g., less than absolute subunit resolution, where "similarity" in the sequence will be detectable, or the form of selective sequences of homology interspersed at specific or irregular locations.

In either case, the sequence is determinable at selective resolution or at particular locations. Thus, the hybridization method will be useful as a means for identification, e.g., a "fingerprint", much like a Southern hybridization method is used. It is also useful to map particular target sequences.

B. VLSIPS™ Technology

The invention is enabled by the development of technology to prepare substrates on which specific reagents may be either positionally attached or synthesized. In particular, the very large scale immobilized polymer synthesis (VLSIPS™) technology allows for the very high density production of an enormous diversity of reagents mapped out in a known matrix pattern on a substrate. These reagents specifically recognize subsequences in a target polymer and bind thereto, producing a map of positionally defined regions of interaction. These map positions are convertible into actual features recognized, and thus would be present in the target molecule of interest.

As indicated, the sequence specific recognition reagents will often be oligonucleotides which hybridize with fidelity and discrimination to the target sequence. For use with other polymers, monoclonal or polyclonal antibodies having high sequence specificity will often be used.

In the generic sense, the VLSIPS technology allows the production of a substrate with a high density matrix of positionally mapped regions with specific recognition reagents attached at each distinct region. By use of protective groups which can be positionally removed, or added, the regions can be activated or deactivated for addition of particular reagents or compounds. Details of the protection are described below and in related Pirrung et al. (1992) U.S. Pat. No. 5,143,854. In a preferred embodiment, photosensitive protecting agents will be used and the regions of activation or deactivation may be controlled by electro-optical and optical methods, similar to many of the processes used in semiconductor wafer and chip fabrication.

In the nucleic acid nucleotide sequencing application, a VLSIPS substrate is synthesized having positionally defined oligonucleotide probes. See Pirrung et al. (1992) U.S. Pat. No. 5,143,854; and U.S. Ser. No. 07/624,120, now abandoned. By use of masking technology and photosensitive synthetic subunits, the VLSIPS apparatus allows for the stepwise synthesis of polymers according to a positionally defined matrix pattern. Each oligonucleotide probe will be synthesized at known and defined positional locations on the substrate. This forms a matrix pattern of known relationship between position and specificity of interaction. The VLSIPS technology allows the production of a very large number of different oligonucleotide probes to be simultaneously and automatically synthesized including numbers in excess of about 10^2, 10^3, 10^4, 10^5, 10^6, or even more, and at densities of at least about 10^2, $10^3/cm^2$, $10^4/cm^2$, $10^5/cm^2$ and up to $10^6/cm^2$ or more. This application discloses methods for synthesizing polymers on a silicon or other suitably derivatized substrate, methods and chemistry for synthesizing specific types of biological polymers on those substrates, apparatus for scanning and detecting whether interaction has occurred at specific locations on the substrate, and various other technologies related to the use of a high density very large scale immobilized polymer substrate. In particular, sequencing, fingerprinting, and mapping applications are discussed herein in detail, though related technologies are described in simultaneously filed applications U.S. Ser. No. 07/624,120, now abandoned; and U.S. Ser. No. 07/517,659; Dower et al. (1995) U.S. Pat. No. 5,427,908, each of which is hereby incorporated herein by reference.

In other embodiments, antibody probes will be generated which specifically recognize particular subsequences found on a polymer. Antibodies would be generated which are specific for recognizing a three contiguous amino acid sequence, and monoclonal antibodies may be preferred. optimally, these antibodies would not recognize any sequences other than the specific three amino acid stretch desired and the binding affinity should be insensitive to flanking or remote sequences found on a target molecule. Likewise, antibodies specific for particular carbohydrate linkages or sequences will be generated. A similar approach could be used for preparing specific reagents which recognize other polymer subunit sequences. These reagents would typically be site specifically localized to a substrate matrix pattern where the regions are closely packed.

These reagents could be individually attached at specific sites on the substrate in a matrix by an automated procedure where the regions are positionally targeted by some other specific mechanism, e.g., one which would allow the entire collection of reagents to be attached to the substrate in a single reaction. Each reagent could be separately attached to a specific oligonucleotide sequence by an automated procedure. This would produce a collection of reagents where, e.g., each monoclonal antibody would have a unique oligonucleotide sequence attached to it. By virtue of a VLSIPS substrate which has different complementary oligonucleotides synthesized on it, each monoclonal antibody would specifically be bound only at that site on the substrate where the complementary oligonucleotide has been synthesized. A crosslinking step would fix the reagent to the substrate. See, e.g., Dattagupta et al. (1985) U.S. Pat. No. 4,542,102 and (1987) U.S. Pat. No. 4,713,326; and Chatterjee, M. et al. (1990) *J. Am. Chem. Soc.* 112:6397–6399, which are hereby incorporated herein by reference. This allows a high density positionally specific collection of specific recognition reagents, e.g., monoclonal antibodies, to be immobilized to a solid substrate using an automated system.

The regions which define particular reagents will usually be generated by selective protecting groups which may be activated or deactivated. Typically the protecting group will be bound to a monomer subunit or spatial region, and can be spatially affected by an activator, such as electromagnetic radiation. Examples of protective groups with utility herein include nitroveratryl oxycarbonyl (NVOC), nitrobenzyl oxycarbony (NBOC), dimethyl dimethoxy benzyloxy carbonyl, 5-bromo-7-nitroindolinyl, O-hydroxy-α-methyl cinnamoyl, and 2-oxymethylene anthraquinone. Examples of activators include ion beams, electric fields, magnetic fields, electron beams, x-ray, and other forms of electromagnetic radiation.

C. Binary Masking

In fact, the means for producing a substrate useful for these techniques are explained in Pirrung et al. (1992) U.S. Pat. No. 5,143,854, which is hereby incorporated herein by reference. However, there are various particular ways to optimize the synthetic processes. Many of these methods are described in Ser. No. 07/624,120, now abandoned.

Briefly, the binary synthesis strategy refers to an ordered strategy for parallel synthesis of diverse polymer sequences by sequential addition of reagents which may be represented by a reactant matrix, and a switch matrix, the product of which is a product matrix. A reactant matrix is a 1×n matrix of the building blocks to be added. The switch matrix is all or a subset of the binary numbers from 1 to n arranged in columns. In preferred embodiments, a binary strategy is one in which at least two successive steps illuminate half of a region of interest on the substrate. In most preferred embodiments, binary synthesis refers to a synthesis strategy which also factors a previous addition step. For example, a strategy in which a switch matrix for a masking strategy halves regions that were previously illuminated, illuminating about half of the previously illuminated region and protecting the remaining half (while also protecting about half of previously protected regions and illuminating about half of previously protected regions). It will be recognized that binary rounds may be interspersed with non-binary rounds and that only a portion of a substrate may be subjected to a binary scheme, but will still be considered to be a binary masking scheme within the definition herein. A binary "masking" strategy is a binary synthesis which uses light to remove protective groups from materials for addition of other materials such as nucleotides or amino acids.

In particular, this procedure provides a simplified and highly efficient method for saturating all possible sequences of a defined length polymer. This masking strategy is also particularly useful in producing all possible oligonucleotide sequence probes of a given length.

D. Applications

The technology provided by the present invention has very broad applications. Although described specifically for polynucleotide sequences, similar sequencing, fingerprinting, mapping, and screening procedures can be applied to polypeptide, carbohydrate, or other polymers. In particular, the present invention may be used to completely sequence a given target sequence to subunit resolution. This may be for de novo sequencing, or may be used in conjunction with a second sequencing procedure to provide independent verification. See, e.g., (1988) *Science* 242:1245. For example, a large polynucleotide sequence defined by either the Maxam and Gilbert technique or by the Sanger technique may be verified by using the present invention.

In addition, by selection of appropriate probes, a polynucleotide sequence can be fingerprinted. Fingerprinting is a less detailed sequence analysis which usually involves the characterization of a sequence by a combination of defined features. Sequence fingerprinting is particularly useful because the repertoire of possible features which can be tested is virtually infinite. Moreover, the stringency of matching is also variable depending upon the application. A Southern Blot analysis may be characterized as a means of simple fingerprint analysis.

Fingerprinting analysis may be performed to the resolution of specific nucleotides, or may be used to determine homologies, most commonly for large segments. In particular, an array of oligonucleotide probes of virtually any workable size may be positionally localized on a matrix and used to probe a sequence for either absolute complementary matching, or homology to the desired level of stringency using selected hybridization conditions.

In addition, the present invention provides means for mapping analysis of a target sequence or sequences. Mapping will usually involve the sequential ordering of a plurality of various sequences, or may involve the localization of a particular sequence within a plurality of sequences. This may be achieved by immobilizing particular large segments onto the matrix and probing with a shorter sequence to determine which of the large sequences contain that smaller sequence. Alternatively, relatively shorter probes of known or random sequence may be immobilized to the matrix and a map of various different target sequences may be determined from overlaps. Principles of such an approach are described in some detail by Evans et al. (1989) "Physical Mapping of Complex Genomes by Cosmid Multiplex Analysis," *Proc. Natl. Acad. Sci. USA* 86:5030–5034; Michiels et al. (1987) "Molecular Approaches to Genome Analysis: A Strategy for the Construction of Ordered Overlap Clone Libraries," *CABIOS* 3:203–210; Olsen et al. (1986) "Random-Clone Strategy for Genomic Restriction Mapping in Yeast," *Proc. Natl. Acad. Sci. USA* 83:7826–7830; Craig, et al. (1990) "Ordering of Cosmid Clones Covering the Herpes Simplex Virus Type I (HSV-I) Genome: A Test Case for Fingerprinting by Hybridization," *Nuc. Acids Res.* 18:2653–2660; and Coulson, et al. (1986) "Toward a Physical Map of the Genome of the Nematode Caenorhabditis elegans," *Proc. Natl. Acad. Sci. USA* 83:7821–7825; each of which is hereby incorporated herein by reference.

Fingerprinting analysis also provides a means of identification. In addition to its value in apprehension of criminals from whom a biological sample, e.g., blood, has been collected, fingerprinting can ensure personal identification for other reasons. For example, it may be useful for identification of bodies in tragedies such as fire, flood, and vehicle crashes. In other cases the identification may be useful in identification of persons suffering from amnesia, or of missing persons. Other forensics applications include establishing the identity of a person, e.g., military identification "dog tags", or may be used in identifying the source of particular biological samples. Fingerprinting technology is described, e.g., in Carrano, et al. (1989) "A High-Resolution, Fluorescence-Based, Semi-automated method for DNA Fingerprinting," *Genomics* 4: 129–136, which is hereby incorporated herein by reference. See, e.g., table I, for nucleic acid applications, and corresponding applications may be accomplished using polypeptides.

TABLE I

VLSIPS ™ TECHNOLOGY IN NUCLEIC ACIDS

- I. Construction of Chips
- II. Applications
  - A. Sequencing
    1. Primary sequencing
    2. Secondary sequencing (sequence checking)
    3. Large scale mapping
    4. Fingerprinting
  - B. Duplex/Triplex formation
    1. Antisense
    2. Sequence specific function modulation (e.g. promoter inhibition)
  - C. Diagnosis
    1. Genetic markers
    2. Type markers
       a. Blood donors
       b. Tissue transplants

TABLE I-continued

VLSIPS ™ TECHNOLOGY IN NUCLEIC ACIDS
D. Microbiology
1. Clinical microbiology
2. Food microbiology
III. Instrumentation
A. Chip machines
B. Detection
IV. Software Development
A. Instrumentation software
B. Data reduction software
C. Sequence analysis software

The fingerprinting analysis may be used to perform various types of genetic screening. For example, a single substrate may be generated with a plurality of screening probes, allowing for the simultaneous genetic screening for a large number of genetic markers. Thus, prenatal or diagnostic screening can be simplified, economized, and made more generally accessible.

In addition to the sequencing, fingerprinting, and mapping applications, the present invention also provides means for determining specificity of interaction with particular sequences. Many of these applications were described in Ser. No. 07/362,901, now abandoned, Pirrung et al. (1992) U.S. Pat. No. 5,143,854; Ser. No. 07/435,316, and Ser. No. 07/612,671.

E. Detection Methods and Apparatus

An appropriate detection method applicable to the selected labeling method can be selected. Suitable labels include radionucleotides, enzymes, substrates, cofactors, inhibitors, magnetic particles, heavy metal atoms, and particularly fluorescers, chemiluminescers, and spectroscopic labels. Patents teaching the use of such labels include U.S. Pat. Nos. 3,817,837; 3,850,752; 3,939,350; 3,996,345; 4,277,437; 4,275,149; and 4,366,241.

With an appropriate label selected, the detection system best adapted for high resolution and high sensitivity detection may be selected. As indicated above, an optically detectable system, e.g., fluorescence or chemiluminescence would be preferred. Other detection systems may be adapted to the purpose, e.g., electron microscopy, scanning electron microscopy (SEM), scanning tunneling electron microscopy (STEM), infrared microscopy, atomic force microscopy (AFM), electrical condutance, and image plate transfer.

With a detection method selected, an apparatus for scanning the substrate will be designed. Apparatus, as described in Ser. No. 07/362,901, now abandoned; or Pirrung et al. (1992) U.S. Pat. No. 5,143,854; or Ser. No. 07/624,120, now abandoned, are particularly appropriate. Design modifications may also be incorporated therein.

F. Data Analysis

Data is analyzed by processes similar to those described below in the section describing theoretical analysis. More efficient algorithms will be mathematically devised, and will usually be designed to be performed on a computer. Various computer programs which may more quickly or efficiently make measurement samples and distinguish signal from noise will also be devised. See, particularly, Ser. No. 07/624,120, now abandoned.

The initial data resulting from the detection system is an array of data indicative of fluorescent intensity versus location on the substrate. The data are typically taken over regions substantially smaller than the area in which synthesis of a given polymer has taken place. Merely by way of example, if polymers were synthesized in squares on the substrate having dimensions of 500 microns by 500 microns, the data may be taken over regions having dimensions of 5 microns by 5 microns. In most preferred embodiments, the regions over which florescence data are taken across the substrate are less than about ½ the area of the regions in which individual polymers are synthesized, preferably less than 1/10 the area in which a single polymer is synthesized, and most preferably less than 1/100 the area in which a single polymer is synthesized. Hence, within any area in which a given polymer has been synthesized, a large number of fluorescence data points are collected.

A plot of number of pixels versus intensity for a scan should bear a rough resemblance to a bell curve, but spurious data are observed, particularly at higher intensities. Since it is desirable to use an average of fluorescent intensity over a given synthesis region in determining relative binding affinity, these spurious data will tend to undesirably skew the data.

Accordingly, in one embodiment of the invention the data are corrected for removal of these spurious data points, and an average of the data points is thereafter utilized in determining relative binding efficiency. In general the data are fitted to a base curve and statistical measures are used to remove spurious data.

In an additional analytical tool, various degeneracy reducing analogues may be incorporated in the hybridization probes. Various aspects of this strategy are described, e.g., in Macevicz, S. (1990) PCT publication number WO 90/04652, which is hereby incorporated herein by reference.

II. THEORETICAL ANALYSIS

The principle of the hybridization sequencing procedure is based, in part, upon the ability to determine overlaps of short segments. The VLSIPS technology provides the ability to generate reagents which will saturate the possible short subsequence recognition possibilities. The principle is most easily illustrated by using a binary sequence, such as a sequence of zeros and ones. Once having illustrated the application to a binary alphabet, the principle may easily be understood to encompass three letter, four letter, five or more letter, even 20 letter alphabets. A theoretical treatment of analysis of subsequence information to reconstruction of a target sequence is provided, e.e., in Lysov, Yu., et al. (1988) *Doklady Akademi. Nauk. SSR* 303:1508–1511; Khrapko K., et al. (1989) *FEBS Letters* 256:118–122; Pevzner, P. (1989) *J. of Biomolecular Structure and Dynamics* 7:63–69; and Drmanac, R. et al. (1989) *Genomics* 4:114–128; each of which is hereby incorporated herein by reference.

The reagents for recognizing the subsequences will usually be specific for recognizing a particular polymer subsequence anywhere within a target polymer. It is preferable that conditions may be devised which allow absolute discrimination between high fidelity matching and very low levels of mismatching. The reagent interaction will preferably exhibit no sensitivity to flanking sequences, to the subsequence position within the target, or to any other remote structure within the sequence. For polynucleotide sequencing, the specific reagents can be oligonucleotide probes; for polypeptides and carbohydrates, antibodies will be useful reagents. Antibody reagents should also be useful for other types of polymers.

A. Simple n-mer Structure: Theory

1. Simple Two Letter Alphabet: Example

A simple example is presented below of how a sequence of ten digits comprising zeros and ones would be sequenceable using short segments of five digits. For example, consider the sample ten digit sequence:

1010011100.

A VLSIPS™ Technology substrate could be constructed, as discussed elsewhere, which would have reagents attached in

a defined matrix pattern which specifically recognize each of the possible five digit sequences of ones and zeros. The number of possible five digit subsequences is 2^5=32. The number of possible different sequences 10 digits long is 2^{10}=1,024. The five contiguous digit subsequences within a ten digit sequence number six, i.e., positioned at digits 1–5, 2–6, 3–7, 4–8, 5–9, and 6–10. It will be noted that the specific order of the digits in the sequence is important and that the order is directional, e.g., running left to right versus right to left. The first five digit sequence contained in the target sequence is 10100. The second is 01001, the third is 10011, the fourth is 00111, the fifth is 01110, and the sixth is 11100.

The VLSIPS™ substrate would have a matrix pattern of positionally attached reagents which recognize each of the different 5-mer subsequences. Those reagents which recognize each of the 6 contained 5-mers will bind the target, and a label allows the positional determination of where the sequence specific interaction has occurred. By correlation of the position in the matrix pattern, the corresponding bound subsequences can be determined.

In the above-mentioned sequence, six different 5-mer sequences would be determined to be present. They would be:

10100

01001

10011

00111

01110

11100

Any sequence which contains the first five digit sequence, 10100, already narrows the number of possible sequences (e.g., from 1024 possible sequences) which contain it to less than about 192 possible sequences.

This 192 is derived from the observation that with the subsequence 10100 at the far left of the sequence, in positions 1–5, there are only 32 possible sequences. Likewise, for that particular subsequence in positions 2–6, 3–7, 4–8, 5–9, and 6–10. So, to sum up all of the sequences that could contain 10100, there are 32 for each position and 6 positions for a total of about 192 possible sequences. However, some of these 10 digit sequences will have been counted twice. Thus, by virtue of containing the 10100 subsequence, the number of possible 10-mer sequences has been decreased from 1024 sequences to less than about 192 sequences.

In this example, not only do we know that the sequence contains 10100, but we also know that it contains the second five character sequence, 01001. By virtue of knowing that the sequence contains 10100, we can look specifically to determine whether the sequence contains a subsequence of five characters which contains the four leftmost digits plus a next digit to the left. For example, we would look for a sequence of X1010, but we find that there is none. Thus, we know that the 10100 must be at the left end of the 10-mer. We would also look to see whether the sequence contains the rightmost four digits plus a next digit to the right, e.g., 0100X. We find that the sequence also contains the sequence 01001, and that X is a 1. Thus, we know at least that our target sequence has an overlap of 0100 and has the left terminal sequence 101001.

Applying the same procedure to the second 5-mer, we also know that the sequence must include a sequence of five digits having the sequence 1001Y where Y must be either 0 or 1. We look through the fragments and we see that we have a 10011 sequence within our target, thus Y is also 1. Thus, we would know that our sequence has a sequence of the first seven being 1010011.

Moving to the next 5-mer, we know that there must be a sequence of 0011Z, where Z must be either 0 or 1. We look at the fragments produced above and see that the target sequence contains a 00111 subsequence and Z is 1. Thus, we know the sequence must start with 10100111.

The next 5-mer must be of the sequence 0111W where W must be 0 or 1. Again, looking up at the fragments produced, we see that the target sequence contains a 01110 subsequence, and W is a 0. Thus, our sequence to this point is 101001110. We know that the last 5-mer must be either 11100 or 11101. Looking above, we see that it is 11100 and that must be the last of our sequence. Thus, we have determined that our sequence must have been 1010011100.

However, it will be recognized from the example above with the sequences provided therein, that the sequence analysis can start with any known positive probe subsequence. The determination may be performed by moving linearly along the sequence checking the known sequence with a limited number of next positions. Given this possibility, the sequence may be determined, besides by scanning all possible oligonucleotide probe positions, by specifically looking only where the next possible positions would be. This may increase the complexity of the scanning but may provide a longer time span dedicated towards scanning and detecting specific positions of interest relative to other sequence possibilities. Thus, the scanning apparatus could be set up to work its way along a sequence from a given contained oligonucleotide to only look at those positions on the substrate which are expected to have a positive signal.

It is seen that given a sequence, it can be de-constructed into n-mers to produce a set of internal contiguous subsequences. From any given target sequence, we would be able to determine what fragments would result. The hybridization sequence method depends, in part, upon being able to work in the reverse, from a set of fragments of known sequences to the full sequence. In simple cases, one is able to start at a single position and work in either or both directions towards the ends of the sequence as illustrated in the example.

The number of possible sequences of a given length increases very quickly with the length of that sequence. Thus, a 10-mer of zeros and ones has 1024 possibilities, a 12-mer has 4096. A 20-mer has over a million possibilities, and a 30-mer has over a billion. However, a given 30-mer has, at most, 26 different internal 5-mer sequences. Thus, a 30 character target sequence having over a million possible sequences can be substantially defined by only 26 different 5-mers. It will be recognized that the probe oligonucleotides will preferably, but need not necessarily, be of identical length, and that the probe sequences need not necessarily be contiguous in that the overlapping subsequences need not differ by only a single subunit. Moreover, each position of the matrix pattern need not be homogeneous, but may actually contain a plurality of probes of known sequence. In addition, although all of the possible subsequence specifications would be preferred, a less than full set of sequences specifications could be used. In particular, although a substantial fraction will preferably be at least about 70%, it may be less than that. About 20% would be preferred, more preferably at least about 30% would be desired. Higher percentages would be especially preferred.

2. Example of Four Letter Alphabet

A four letter alphabet may be conceptualized in at least two different ways from the two letter alphabet. One way is to consider the four possible values at each position and to analogize in a similar fashion to the binary example each of the overlaps. A second way is to group the binary digits into groups.

Using the first means, the overlap comparisons are performed with a four letter alphabet rather than a two letter alphabet. Then, in contrast to the binary system with 10 positions where 2^{10}=1024 possible sequences, in a 4-character alphabet with 10 positions, there will actually be 4^{10}=1,048,576 possible sequences. Thus, the complexity of a four character sequence has a much larger number of possible sequences compared to a two character sequence. Note, however, that there are still only 6 different internal 5-mers. For simplicity, we shall examine a 5 character string with 3 character subsequences. Instead of only 1 and 0, the characters may be designated, e.g., A, C, G, and T. Let us take the sequence GGCTA. The 3-mer subsequences are:

GGC
GCT
CTA

Given these subsequences, there is one sequence, or at most only a few sequences which would produce that combination of subsequences, i.e., GGCTA.

Alternatively, with a four character universe, the binary system can be looked at in pairs of digits. The pairs would be 00, 01, 10, and 11. In this manner, the earlier used sequence 1010011100 is looked at as 10,10,01,11,00. Then the first character of two digits is selected from the possible universe of the four representations 00, 01, 10, and 11. Then a probe would be in an even number of digits, e.g., not five digits, but, three pairs of digits or six digits. A similar comparison is performed and the possible overlaps determined. The 3-pair subsequences are:

10, 10, 01
10, 01, 11
01, 11, 00

and the overlap reconstruction produces 10,10,01,11,00.

The latter of the two conceptual views of the 4 letter alphabet provides a representation which is similar to what would be provided in a digital computer. The applicability to a four nucleotide alphabet is easily seen by assigning, e.g., 00 to A, 01 to C, 10 to G, and 11 to T. And, in fact, if such a correspondence is used, both examples for the 4 character sequences can be seen to represent the same target sequence. The applicability of the hybridization method and its analysis for determining the ultimate sequence is easily seen if A is the representation of adenine, C is the representation of cytosine, G is the representation of guanine, and T is the representation of thymine or uracil.

3. Generalization to m-letter Alphabet

This reconstruction process may be applied to polymers of virtually any number of possible characters in the alphabet, and for virtually any length sequence to be sequenced, though limitations, as discussed below, will limit its efficiency at various extremes of length. It will be recognized that the theory can be applied to a large diversity of systems where sequence is important.

For example, the method could be applied to sequencing of a polypeptide. A polypeptide can have any of twenty natural amino acid possibilities at each position. A twenty letter alphabet is amenable to sequencing by this method so long as reagents exist for recognizing shorter subsequences therein. A preferred reagent for achieving that goal would be a set of monoclonal antibodies each of which recognizes a specific three contiguous amino acid subsequence. A complete set of antibodies which recognize all possible subsequences of a given length, e.g., 3 amino acids, and preferably with a uniform affinity, would be 20^3=8000 reagents.

It will also be recognized that each target sequence which is recognized by the specific reagents need not have homogeneous termini. Thus, fragments of the entire target sequence will also be useful for hybridizing appropriate subsequences. It is, however, preferable that there not be a significant amount of labeled homogeneous contaminating extraneous sequences. This constraint does usually require the purification of the target molecule to be sequenced, but a specific label technique would dispense with a purification requirement if the unlabeled extraneous sequences do not interfere with the labeled sequences.

In addition, conformational effects of target polypeptide folding may, in certain embodiments, be negligible if the polypeptide is fragmented into sufficiently small peptides, or if the interaction is performed under conditions where conformation, but not specific interaction, is disrupted.

B. Complications

Two obvious complications exist with the method of sequence analysis by hybridization. The first results from a probe of inappropriate length while the second relates to internally repeated sequences.

The first obvious complication is a problem which arises from an inappropriate length of recognition sequence, which causes problems with the specificity of recognition. For example, if the recognized sequence is too short, every sequence which is utilized will be recognized by every probe sequence. This occurs, e.g., in a binary system where the probes are each of sequences which occur relatively frequently, e.g., a two character probe for the binary system. Each possible two character probe would be expected to appear ¼ of the time in every single two character position. Thus, the above sequence example would be recognized by each of the 00, 10, 01, and 11. Thus, the sequence information is virtually lost because the resolution is too low and each recognition reagent specifically binds at multiple sites on the target sequence.

The number of different probes which bind to a target depends on the relationship between the probe length and the target length. At the extreme of short probe length, the just mentioned problem exists of excessive redundancy and lack of resolution. The lack of stability in recognition will also be a problem with extremely short probes. At the extreme of long probe length, each entire probe sequence is on a different position of a substrate. However, a problem arises from the number of possible sequences, which goes up dramatically with the length of the sequence. Also, the specificity of recognition begins to decrease as the contribution to binding by any particular subunit may become sufficiently low that the system fails to distinguish the fidelity of recognition. Mismatched hybridization may be a problem with the polynucleotide sequencing applications, though the fingerprinting and mapping applications may not be so strict in their fidelity requirements. As indicated above, a thirty position binary sequence has over a million possible sequences, a number which starts to become unreasonably large in its required number of different sequences, even

though the target length is still very short. Preparing a substrate with all sequence possibilities for a long target may be extremely difficult due to the many different oligomers which must be synthesized.

The above example illustrates how a long target sequence may be reconstructed with a reasonably small number of shorter subsequences. Since the present day resolution of the regions of the substrate having defined oligomer probes attached to the substrate approaches about 10 microns by 10 microns for resolvable regions, about 10^6, or 1 million, positions can be placed on a one centimeter square substrate. However, high resolution systems may have particular disadvantages which may be outweighed using the lower density substrate matrix pattern. For this reason, a sufficiently large number of probe sequences can be utilized so that any given target sequence may be determined by hybridization to a relatively small number of probes.

A second complication relates to convergence of sequences to a single subsequence. This will occur when a particular subsequence is repeated in the target sequence. This problem can be addressed in at least two different ways. The first, and simpler way, is to separate the repeat sequences onto two different targets. Thus, each single target will not have the repeated sequence and can be analyzed to its end. This solution, however, complicates the analysis by requiring that some means for cutting at a site between the repeats can be located. Typically a careful sequencer would want to have two intermediate cut points so that the intermediate region can also be sequenced in both directions across each of the cut points. This problem is inherent in the hybridization method for sequencing but can be minimized by using a longer known probe sequence so that the frequency of probe repeats is decreased.

Knowing the sequence of flanking sequences of the repeat will simplify the use of polymerase chain reaction (PCR) or a similar technique to further definitively determine the sequence between sequence repeats. Probes can be made to hybridize to those known sequences adjacent the repeat sequences, thereby producing new target sequences for analysis. See, e.g., Innis et al. (eds.) (1990) *PCR Protocols: A Guide to Methods and Applications, Academic Press*; and methods for synthesis of oligonucleotide probes, see, e.g., Gait (1984) *Oligonucleotide Synthesis: A Practical Approach*, IRL Press, Oxford.

Other means for dealing with convergence problems include using particular longer probes, and using degeneracy reducing analogues, see, e.g., Macevicz, S. (1990) PCT publication number WO 90/04652, which is hereby incorporated herein by reference. By use of stretches of the degeneracy reducing analogues with other probes in particular combinations, the number of probes necessary to fully saturate the possible oligomer probes is decreased. For example, with a stretch of 12-mers having the central 4-mer of degenerate nucleotides, in combination with all of the possible 8-mers, the collection numbers twice the number of possible 8-mers, e.g. 65,536+65,536=131,072, but the population provides screening equivalent to all possible 12-mers.

By way of further explanation, all possible oligonucleotide 8-mers may be depicted in the fashion:

N1-N2-N3-N4-N5-N6-N7-N8,

in which there are 4^8=65,536 possible 8-mers. As described in Ser. No. 07/624,120, now abandoned, producing all possible 8-mers requires 4×8=32 chemical binary synthesis steps to produce the entire matrix pattern of 65,536 8-mer possibilities. By incorporating degeneracy reducing nucleotides, D's, which hybridize nonselectively to any corresponding complementary nucleotide, new oligonucleotides 12-mers can be made in the fashion:

N1-N2-N3-N4-D-D-D-D-N5-N6-N7-N8,

in which there are again, as above, only 4^8=65,536 possible "12-mers", which in reality only have 8 different nucleotides.

However, it can be seen that each possible 12-mer probe could be represented by a group of the two 8-mer types. Moreover, repeats of less than 12 nucleotides would not converge, or cause repeat problems in the analysis. Thus, instead of requiring a collection of probes corresponding to all 12-mers, or 4^{12}=16,777,216 different 12-mers, the same information can be derived by making 2 sets of "8-mers" consisting of the typical 8-mer collection of 4^8=65,536 and the "12-mer" set with the degeneracy reducing analogues, also requiring making 4^8=65,536. The combination of the two sets, requires making 65,536+65,536=131,072 different molecules, but giving the information of 16,777,216 molecules. Thus, incorporating the degeneracy reducing analogue decreases the number of molecules necessary to get 12-mer resolution by a factor of about 128-fold.

C. Non-polynucleotide Embodiments

The above example is directed towards a polynucleotide embodiment. This application is relatively easily achieved because the specific reagents will typically be complementary oligonucleotides, although in certain embodiments other specific reagents may be desired. For example, there may be circumstances where other than complementary base pairing will be utilized. The polynucleotide targets, will usually be single strand, but may be double or triple stranded in various applications. However, a triple stranded specific interaction might be sometimes desired, or a protein or other specific binding molecule may be utilized. For example, various promoter or DNA sequence specific binding proteins might be used, including, e.g., restriction enzyme binding domains, other binding domains, and antibodies. Thus, specific recognition reagents besides oligonucleotides may be utilized.

For other polymer targets, the specific reagents will often be polypeptides. These polypeptides may be protein binding domains from enzymes or other proteins which display specificity for binding. Usually an antibody molecule may be used, and monoclonal antibodies may be particularly desired. Classical methods may be applied for preparing antibodies, see, e.g., Harlow and Lane (1988) *Antibodies: A Laboratory Manual* Cold Spring Harbor Press, New York; and Goding (1986) *Monoclonal Antibodies: Principles and Practice* (2d Ed.) Academic Press, San Diego. Other suitable techniques for in vitro exposure of lymphocytes to the antigens or selection of libraries of antibody binding sites are described, e.g., in Huse et al. (1989) *Science* 246:1275–1281; and Ward et al. 91989) *Nature* 341:544–546, each of which is hereby incorporated herein by reference. Unusual antibody production methods are also described, e.g., in Hendricks et al. (1989) *BioTechnology* 7:1271–1274; and Hiatt et al. (1989) *Nature* 342:76–78, each of which is hereby incorporated herein by reference. Other molecules which may exhibit specific binding interaction may be useful for attachment to a VLSIPS substrate by various methods, including the caged biotin methods, see, e.g., Ser. No. 07/435,316, now abandoned, and Barrett et al. (1993) U.S. Pat. No. 5,252,743.

The antibody specific reagents should be particularly useful for the polypeptide, carbohydrate, and synthetic poly-

mer applications. Individual specific reagents might be generated by an automated process to generate the number of reagents necessary to advantageously use the high density positional matrix pattern. In an alternative approach, a plurality of hybridoma cells may be screened for their ability to bind to a VLSIPS matrix possessing the desired sequences whose binding specificity is desired. Each cell might be individually grown up and its binding specificity determined by VLSIPS apparatus and technology. An alternative strategy would be to expose the same VLSIPS matrix to a polyclonal serum of high titer. By a successively large volume of serum and different animals, each region of the VLSIPS substrate would have attached to it a substantial number of antibody molecules with specificity of binding. The substrate, with non-covalently bound antibodies could be derivatized and the antibodies transferred to an adjacent second substrate in the matrix pattern in which the antibody molecules had attached to the first matrix. If the sensitivity of detection of binding interaction is sufficiently high, such a low efficiency transfer of antibody molecules may produce a sufficiently high signal to be useful for many purposes, including the sequencing applications.

In another embodiment, capillary forces may be used to transfer the selected reagents to a new matrix, to which the reagents would be positionally attached in the pattern of the recognized sequences. Or, the reagents could be transversely electrophoresed, magnetically transferred, or otherwise transported to a new substrate in their retained positional pattern.

III. POLYNUCLEOTIDE SEQUENCING

In principle, the making of a substrate having a positionally defined matrix pattern of all possible oligonucleotides of a given length involves a conceptually simple method of synthesizing each and every different possible oligonucleotide, and affixing them to a definable position. Oligonucleotide synthesis is presently mechanized and enabled by current technology, see, e.g., Ser. No. 07/362,901, now abandoned; Pirrung et al. (1992) U.S. Pat. No. 5,143,854; and instruments supplied by Applied Biosystems, Foster City, Calif.

A. Preparation of Substrate Matrix

The production of the collection of specific oligonucleotides used in polynucleotide sequencing may be produced in at least two different ways. Present technology certainly allows production of ten nucleotide oligomers on a solid phase or other synthesizing system. See, e.g., instrumentation provided by Applied Biosystems, Foster City, Calif. Although a single oligonucleotide can be relatively easily made, a large collection of them would typically require a fairly large amount of time and investment. For example, there are 4^{10}=1,048,576 possible ten nucleotide oligomers. Present technology allows making each and every one of them in a separate purified form though such might be costly and laborious.

Once the desired repertoire of possible oligomer sequences of a given length have been synthesized, this collection of reagents may be individually positionally attached to a substrate, thereby allowing a batchwise hybridization step. Present technology also would allow the possibility of attaching each and every one of these 10-mers to a separate specific position on a solid matrix. This attachment could be automated in any of a number of ways, particularly through the use of a caged biotin type linking. This would produce a matrix having each of different possible 10-mers.

A batchwise hybridization is much preferred because of its reproducibility and simplicity. An automated process of attaching various reagents to positionally defined sites on a substrate is provided in Pirrung et al. (1992) U.S. Pat. No. 5,143,854; Ser. No. 07/624,120, now abandoned; and Barrett et al. (1993) U.S. Pat. No. 5,252,743; each of which is hereby incorporated herein by reference.

Instead of separate synthesis of each oligonucleotide, these oligonucleotides are conveniently synthesized in parallel by sequential synthetic processes on a defined matrix pattern as provided in Pirrung et al. (1992) U.S. Pat. No. 5,143,854; and Ser. No. 07/624,120, now abandoned, which are incorporated herein by reference. Here, the oligonucleotides are synthesized stepwise on a substrate at positionally separate and defined positions. Use of photosensitive blocking reagents allows for defined sequences of synthetic steps over the surface of a matrix pattern. By use of the binary masking strategy, the surface of the substrate can be positioned to generate a desired pattern of regions, each having a defined sequence oligonucleotide synthesized and immobilized thereto.

Although the prior art technology can be used to generate the desired repertoire of oligonucleotide probes, an efficient and cost effective means would be to use the VLSIPS technology described in Pirrung et al. (1992) U.S. Pat. No. 5,143,854 and Ser. No. 07/624,120, now abandoned. In this embodiment, the photosensitive reagents involved in the production of such a matrix are described below.

The regions for synthesis may be very small, usually less than about 100 μm×100 μm, more usually less than about 50 μm×50 μm. The photolithography technology allows synthetic regions of less than about 10 μm×10 μm, about 3 μm×3 μm, or less. The detection also may detect such sized regions, though larger areas are more easily and reliably measured.

At a size of about 30 microns by 30 microns, one million regions would take about 11 centimeters square or a single wafer of about 4 centimeters by 4 centimeters. Thus the present technology provides for making a single matrix of that size having all one million plus possible oligonucleotides. Region size is sufficiently small to correspond to densities of at least about 5 regions/cm^2, 20 regions/cm^2, 50 regions/cm^2, 100 regions/cm^2, and greater, including 300 regions/cm^2, 1000 regions/cm^2, 3K regions/cm^2, 10K regions/cm^2, 30K regions/cm^2, 100K regions/cm^2, 300K regions/cm^2 or more, even in excess of one million regions/cm^2.

Although the pattern of the regions which contain specific sequences is theoretically not important, for practical reasons certain patterns will be preferred in synthesizing the oligonucleotides. The application of binary masking algorithms for generating the pattern of known oligonucleotide probes is described in related Ser. No. 07/624,120, now abandoned, which was filed simultaneously with this application. By use of these binary masks, a highly efficient means is provided for producing the substrate with the desired matrix pattern of different sequences. Although the binary masking strategy allows for the synthesis of all lengths of polymers, the strategy may be easily modified to provide only polymers of a given length. This is achieved by omitting steps where a subunit is not attached.

The strategy for generating a specific pattern may take any of a number of different approaches. These approaches are well described in related application Ser. No. 07/624,120, now abandoned, and include a number of binary masking approaches which will not be exhaustively discussed herein. However, the binary masking and binary synthesis approaches provide a maximum of diversity with a minimum number of actual synthetic steps.

The length of oligonucleotides used in sequencing applications will be selected on criteria determined to some extent by the practical limits discussed above. For example, if probes are made as oligonucleotides, there will be 65,536 possible eight nucleotide sequences. If a nine subunit oligonucleotide is selected, there are 262,144 possible permeations of sequences. If a ten-mer oligonucleotide is selected, there are 1,048,576 possible permeations of sequences. As the number gets larger, the required number of positionally defined subunits necessary to saturate the possibilities also increases. With respect to hybridization conditions, the length of the matching necessary to confer stability of the conditions selected can be compensated for. See, e.g., Kanehisa, M. (1984) *Nuc. Acids Res.* 12:203–213, which is hereby incorporated herein by reference.

Although not described in detail here, but below for oligonucleotide probes, the VLSIPS technology would typically use a photosensitive protective group on an oligonucleotide. Sample oligonucleotides are shown in FIG. 1. In particular, the photoprotective group on the nucleotide molecules may be selected from a wide variety of positive light reactive groups preferably including nitro aromatic compounds such as o-nitro-benzyl derivatives or benzylsulfonyl. See, e.g., Gait (1984) *Oligonucleotide Synthesis: A Practical Approach*, IRL Press, Oxford, which is hereby incorporated herein by reference. In a preferred embodiment, 6-nitroveratryl oxycarbony (NVOC), 2-nitrobenzyl oxycarbonyl (NBOC), or α,α-dimethyl-dimethoxybenzyl oxycarbonyl (DEZ) is used. Photoremovable protective groups are described in, e.g., Patchornik (1970) *J. Amer. Chem. Soc.* 92:6333–6335; and Amit et al. (1974) *J. Organic Chem.* 39:192–196; each of which is hereby incorporated herein by reference.

A preferred linker for attaching the oligonucleotide to a silicon matrix is illustrated in FIG. 2. A more detailed description is provided below. A photosensitive blocked nucleotide may be attached to specific locations of unblocked prior cycles of attachments on the substrate and can be successively built up to the correct length oligonucleotide probe.

It should be noted that multiple substrates may be simultaneously exposed to a single target sequence where each substrate is a duplicate of one another or where, in combination, multiple substrates together provide the complete or desired subset of possible subsequences. This provides the opportunity to overcome a limitation of the density of positions on a single substrate by using multiple substrates. In the extreme case, each probe might be attached to a single bead or substrate and the beads sorted by whether there is a binding interaction. Those beads which do bind might be encoded to indicate the subsequence specificity of reagents attached thereto.

Then, the target may be bound to the whole collection of beads and those beads that have appropriate specific reagents on them will bind to the target. Then a sorting system may be utilized to sort those beads that actually bind the target from those that do not. This may be accomplished by presently available cell sorting devices or a similar apparatus. After the relatively small number of beads which have bound the target have been collected, the encoding scheme may be read off to determine the specificity of the reagent on the bead. An encoding system may include a magnetic system, a shape encoding system, a color encoding system, or a combination of any of these, or any other encoding system. Once again, with the collection of specific interactions that have occurred, the binding may be analyzed for sequence information, fingerprint information, or mapping information.

The parameters of polynucleotide sizes of both the probes and target sequences are determined by the applications and other circumstances. The length of the oligonucleotide probes used will depend in part upon the limitations of the VLSIPS technology to provide the number of desired probes. For example, in an absolute sequencing application, it is often useful to have virtually all of the possible oligonucleotides of a given length. As indicated above, there are 65,536 8-mers, 262,144 9-mers, 1,048,576 10-mers, 4,194,304 11-mers, etc. As the length of the oligomer increases the number of different probes which must be synthesized also increases at a rate of a factor of 4 for every additional nucleotide. Eventually the size of the matrix and the limitations in the resolution of regions in the matrix will reach the point where an increase in number of probes becomes disadvantageous. However, this sequencing procedure requires that the system be able to distinguish, by appropriate selection of hybridization and washing conditions, between binding of absolute fidelity and binding of complementary sequences containing mismatches. On the other hand, if the fidelity is unnecessary, this discrimination is also unnecessary and a significantly longer probe may be used. Significantly longer probes would typically be useful in fingerprinting or mapping applications.

The length of the probe is selected for a length that will allow the probe to bind with specificity to possible targets. The hybridization conditions are also very important in that they will determine how closely the homology of complementary binding will be detected. In fact, a single target may be evaluated at a number of different conditions to determine its spectrum of specificity for binding particular probes. This may find use in a number of other applications besides the polynucleotide sequencing fingerprinting or mapping. For example, it will be desired to determine the spectrum of binding affinities and specificities of cell surface antigens with binding by particular antibodies immobilized on the substrate surface, particularly under different interaction conditions. In a related fashion, different regions with reagents having differing affinities or levels of specificity may allow such a spectrum to be defined using a single incubation, where various regions, at a given hybridization condition, show the binding affinity. For example, fingerprint probes of various lengths, or with specific defined non-matches may be used. Unnatural nucleotides or nucleotides exhibiting modified specificity of complementary binding are described in greater detail in Macevicz (1990) PCT pub. No. WO 90/04652; and see the section on modified nucleotides in the Sigma Chemical Company catalogue.

B. Labeling Target Nucleotide

The label used to detect the target sequences will be determined, in part, by the detection methods being applied. Thus, the labeling method and label used are selected in combination with the actual detecting systems being used.

Once a particular label has been selected, appropriate labeling protocols will be applied, as described below for specific embodiments. Standard labeling protocols for nucleic acids are described, e.g., in Sambrook et al.; Kambara, H. et al. (1988) *BioTechnology* 6:816–821; Smith, L. et al. (1985) *Nuc. Acids Res.* 13:2399–2412; for polypeptides, see, e.g., Allen G. (1989) *Sequencing of Proteins and Peptides*, Elsevier, New York, especially chapter 5, and Greenstein and Winitz (1961) *Chemistry of the Amino Acids*, Wiley and Sons, New York. Carbohydrate labeling is described, e.g., in Chaplin and Kennedy (1986) *Carbohydrate Analysis: A Practical Approach*, IRL Press, Oxford. Labeling of other polymers will be performed by methods applicable to them as recognized by a person having ordinary skill in manipulating the corresponding polymer.